Alexander Kokhanovsky   *Editor*

Springer Series in Light Scattering

Volume 3: Radiative Transfer and Light Scattering

Springer

Springer Series in Light Scattering

Series editor

Alexander Kokhanovsky, Vitrociset Belgium, Darmstadt, Germany

The main purpose of new **SPRINGER Series in Light Scattering** is to present recent advances and progress in light scattering media optics. The topic is very broad and incorporates such diverse areas as atmospheric optics, ocean optics, optics of close-packed media, radiative transfer, light scattering, absorption, and scattering by single scatterers and also by systems of particles, biomedical optics, optical properties of cosmic dust, remote sensing of atmosphere and ocean, etc. The topic is of importance for material science, environmental science, climate change, and also for optical engineering. Although main developments in the solutions of radiative transfer and light scattering problems have been achieved in the 20th century by efforts of many scientists including V. Ambartsumian, S. Chandrasekhar, P. Debye, H. C. van de Hulst, G. Mie, and V. Sobolev, the light scattering media optics still have many puzzles to be solved such as radiative transfer in closely packed media, 3D radiative transfer as applied to the solution of inverse problems, optics of terrestrial and planetary surfaces, etc. Also it has a broad range of applications in many brunches of modern science and technology such as biomedical optics, atmospheric and oceanic optics, and astrophysics, to name a few. It is planned that the Series will raise novel scientific questions, integrate data analysis, and offer new insights in optics of light scattering media.

Alexander Kokhanovsky

Editor

Springer Series in Light Scattering

Volume 3: Radiative Transfer and Light Scattering

Springer

Editor
Alexander Kokhanovsky
Vitrociset Belgium
Darmstadt, Hessen
Germany

Contents

Contributors

Matt Christi Fort Collins, Colorado, CO, USA

Anthony B. Davis Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Boaz Ilan Applied Mathematics Department, University of California, Merced, Merced, CA, USA

Olga V. Kalashnikova Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Arnold D. Kim Applied Mathematics Department, University of California, Merced, Merced, CA, USA

Margarita G. Kuzmina Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, Russia

Chao Liu School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, China

Robert Spurr RT Solutions, Inc., Cambridge, MA, USA

The LIDORT and VLIDORT Linearized Scalar and Vector Discrete Ordinate Radiative Transfer Models: Updates in the Last 10 Years

Robert Spurr and Matt Christi

1 Introduction

1.1 General Background

It has been 10 years since the last major review paper on the LIDORT and VLIDORT radiative transfer models; this paper appeared in Light Scattering Reviews, Volume 3 (Spurr (2008), hereinafter referenced as [R1]). The present work will review the current status of the two models in the light of new developments and upgrades in the intervening period. LIDORT stands for "LInearized Discrete Ordinate Radiative Transfer", and the LIDORT model is a comprehensive linearized scalar (no polarization) RT package based on the discrete-ordinate method for solution of the radiative transfer equation (RTE). VLIDORT is the corresponding vector version of LIDORT, dealing with polarized light fields.

Radiative Transfer (RT) models have been around for more than a century, from early investigations in stellar atmospheres, through the pioneering work by Chandrasekhar in the 1940s on the Rayleigh-atmosphere planetary problem, to the wealth of numerical code packages available today. Interestingly, it is only in recent years (see for example Mishchenko 2014 and references therein), that a rigorous derivation of the radiative transfer equation (RTE) has been established for randomly-oriented particles and particle groups, starting from a set of fundamental assumptions in classical electromagnetic scattering theory.

In his famous book on radiative transfer, Chandrasekhar laid out his analytic solution to the Stokes-vector polarized RTE for a single-layer Rayleigh scattering

R. Spurr (✉)
RT Solutions, Inc., 9 Channing Street, Cambridge, MA 02138, USA
e-mail: rtsolutions@verizon.net

M. Christi
Fort Collins, Colorado, CO 80524, USA

medium (Chandrasekhar 1960). Part of the heritage for LIDORT comes from the older DISORT model (Stamnes et al. 1988), which was made available in the late 1980s. DISORT is the first multiple-scattering multi-layer discrete ordinate scalar RT numerical model able to deal with both solar and thermal sources.

The LIDORT model development was started in 1999 in response to a need for a multiple-scattering RT model that would not only calculate radiance fields, but also generate analytically-derived accurate fields of radiance partial derivatives (Jacobians, also known as weighting functions or sensitivity functions) with respect to any atmospheric and/or surface parameters that characterize the optical property inputs to the model. Simultaneous output of calculated intensities and Jacobians is essential for the forward-model component of remote-sensing retrievals based on inverse techniques using functional minimization (with and without regularization) (Rodgers 2000). Jacobians are also important for linear error analyses and sensitivity studies.

RT models having this capability are said to be "linearized". Indeed, the use of linearized RT models is obviously superior to the older technique of finite-difference (FD) Jacobian computation, whereby two separate calls to the RT model are needed to develop a radiance difference to be divided by a perturbation of the desired Jacobian parameter. For applications requiring multi-layer profile Jacobians, FD estimates are slow and cumbersome, requiring many RT model calls; there are also questions of accuracy, since it is a matter of trial and error choosing the degree of perturbation. With LIDORT, a single call to the model will deliver any number of Jacobians of all kinds, and there are no issues of perturbation accuracy.

The LIDORT linearization method is based on end-to-end analytic differentiation of the entire discrete ordinate radiation field in a multi-layer atmosphere, starting with individual layer solutions of the RTE, and moving through the solutions to the boundary-value problem and the post-processing of the solutions using the source-function integration technique (Spurr et al. 2001; Spurr 2002). Other linearized RT models have used the "forward-adjoint" RT solution method to obtain accurate derivatives (Hasekamp and Landgraf 2002; Ustinov 2005; Rozanov and Rozanov 2007; Doicu and Trautmann 2009).

1.2   Historical Overview of the (V)LIDORT Models

There have been a number of versions of LIDORT since its inception in 1999, and similarly for VLIDORT (original version, 2004). Table 1 gives an overview of the model features and associated version numbers. The first LSR paper [R1] covered versions of LIDORT up to and including 3.3, and VLIDORT up to Version 2.4 (orange shading); up to that point, codes were written in Fortran 77. Thereafter, codes were translated to Fortran 90. More details on the historical development of these models may be found in the product User Guides.

Table 1 Major capabilities of LIDORT and VLIDORT

Feature	LIDORT version	VLIDORT version
Pseudo-spherical (solar beam attenuation)	2.1	1.0
Generalized output of Jacobians	2.1	2.0
Green's function treatment	2.3	n/a
3-kernel BRDF system + linearizations	2.4	2.2
Performance (multiple SZA, BVP telescoping)	3.0	2.2
Outgoing sphericity correction	3.2	2.3
Total column Jacobian facility	3.3	2.4
Transmittance-only thermal mode	3.3	2.4
Fortran 90 releases, use of I/O type structures	3.5	2.5
Development of BRDF supplement	3.5	2.5
Development of new surface-leaving supplement	3.6	2.6
Atmospheric and surface Planck-function Jacobians	3.7	2.7
Codes made thread-safe for OpenMP parallel computing	3.7	2.7
Introduction of Taylor series expansions	3.7	2.7
Use of phase functions (matrices) in FO code	3.8	2.8
Do-loop optimizations; bookkeeping improvements	3.8	2.8
BRDF and surface-leaving supplement upgrades	3.8	2.8

Fortran 77 Versions. The first version of LIDORT was developed in 1999 with the profile-Jacobian linearization in a plane-parallel atmosphere (Spurr et al. 2001). Output of weighting functions was limited to the TOA (top-of-atmosphere) upwelling radiation field, with surface property Jacobians only for the Lambertian albedo. Second versions of LIDORT included development of the pseudo-spherical treatment of the solar beam attenuation in a curved atmosphere, and extensions for both upwelling and downwelling Jacobian output at any layer boundary or intermediate level. Green's function methods were introduced for solving the radiative transfer equation (RTE) for solar beam source terms (Spurr 2002; Van Oss and Spurr 2002), and a fully-linearized 3-kernel BRDF (bidirectional reflection distribution function) was developed for the surface boundary condition (Spurr 2004). For the third LIDORT versions, the generation of total column (as opposed to profile) Jacobians was integrated into the code, along with an exact treatment of single scattering for curved line-of-sight paths. Developments for VLIDORT have followed a similar path, starting in 2004 with the first version (no linearization), with second versions developed subsequently following their counterparts with the scalar code (Spurr 2006). All these developments were summed up in the previous review article [R1].

Fortran 90 Versions. In 2010, the LIDORT code was translated to Fortran 90 (Version 3.5), followed by VLIDORT (Version 2.5). The models' organization and coding were overhauled in order to bring the codes in line with modern computing standards.

An important consideration has been the need for the codes to function in a parallel computing environment; this has meant that all COMMON blocks and associated "include" files (prominent features of the F77 codes) were removed, to be replaced by explicit argument declarations for all inputs and outputs, and a number of I/O type structures. Since this translation, the two codes have had equivalent capabilities, and further developments (separation of the BRDF function into a stand-alone supplement, development of similar supplements for surface-leaving capabilities and generation for phase function inputs, generation of new Jacobian facilities such as those for Planck functions, additional performance enhancements including thread-safe code for OpenMP parallel-computing environments) have proceeded in tandem for both codes.

1.3   Scope of this Chapter

The earlier review [R1] contained a full mathematical description of the discrete ordinate theory behind LIDORT and VLIDORT, including a comprehensive treatment of the linearization, both with respect to atmospheric and surface parameters. In the present work, the major aim of the paper is to highlight and describe new developments to the models in the last 10 years, and the mathematical descriptions will be more selective.

Section 2 contains a digest of the discrete-ordinate radiative transfer formulation, and here we develop the RTE solutions and summarize the determination of the complete radiation field and associated atmospheric and surface Jacobians. The emphasis here will be on the scalar treatment in the LIDORT model. Most of the mathematical details will be confined to two appendices.

Section 3 will discuss the LIDORT and VLIDORT model input and output, with special emphasis on the creation of optical property inputs required for Jacobian output; the latter includes new developments for temperature Jacobians in the thermal regime, and "levels-to-layers" conversions for profile Jacobian output. Also in this section we summarize refinements and performance enhancements to the RT solution method (pseudo-spherical approximation, delta-M scaling, N-T ansatz, solution saving and BVP telescoping).

Section 4 deals with the model supplements. One of the main changes since [R1] has been in the treatment of BRDF surfaces—scalar and vector surface reflectance quantities are now generated in separate stand-alone "BRDF supplements" to the main codes. An appendix contains details of the upgraded list of BRDF kernels used in (V)LIDORT. Also, a new supplement has been created for surface-leaving radiation sources (ocean water-leaving, land-based solar-induced fluorescence). Another development is with the FO (first-order) codes, that is, the "exact" single-scatter and direct-thermal parts of the models—these parts have been entirely re-written for the latest releases, and made into self-standing supplemental codes that can either operate on their own or within the greater (V)LIDORT environments; new optical property supplements have been written for the FO models.

With the large and growing community of LIDORT and VLIDORT users, we have included a section on software aspects of the model. Section 5 starts with an overview of the model directory structure, and moves on to benchmarking and performance enhancements—the latter developments in particular include a number of new implementations, including OpenMP parallelization. Section 6 has the concluding remarks.

Other radiative transfer models available from RT Solutions will not be treated here, but they are mentioned in Sect. 5 as part of the benchmarking. In particular, the scalar LIDORT-RRS model (LRRS) (Spurr et al. 2008), which has first-order redistributed rotational-Raman scattering sources, has recently been upgraded and given an analytic Jacobian capability (Spurr and Christi 2018, in preparation). RT Solutions also has a 2-stream model (Spurr and Natraj 2011), and linearized electromagnetic Mie and T-matrix scattering codes (Spurr et al. 2012).

Final note. The present work summarizes the two main models in their current incarnations (Version 2.8 for VLIDORT, Version 3.8 for LIDORT). These versions were completed in 2018, and are publicly available.

2 The LIDORT and VLIDORT Models

In this section, we provide a selective description of the discrete-ordinate formalism at the foundation of the two RT models. Most of the material here is well known from the literature, and a more comprehensive mathematical description may be found in [R1] for instance.

Although Sect. 2.1 introduces the main *vector* RTE, we focus thereafter on the *scalar* LIDORT model and its use of infinite-medium Green's function solution methods for the RTE particular integral. [Additional material on VLIDORT vector RTE solutions will be found in Appendix A as noted in the text.] Section 2.2 deals with determination of discrete-ordinate solutions for the (scalar) homogeneous equations and both particular integrals (solar and thermal sources). Section 2.3 is concerned with the boundary value problem for the whole-atmosphere discrete-ordinate field, and the use of source function integration methods to determine the radiation field at arbitrary levels in the atmosphere and for any viewing geometry. Section 2.4 contains aspects of the linearization process, focusing only on the main results.

2.1 Summary of the Vector RTE

The 1-D vector RTE for plane-parallel scattering in a single layer is:

$$\mu \frac{\partial}{\partial x} \mathbf{I}(x, \mu, \varphi) = -\mathbf{I}(x, \mu, \varphi) + \mathbf{S}(x, \mu, \varphi); \qquad (2.1)$$

$$\mathbf{S}(x, \mu, \varphi) = \frac{\omega(x)}{4\pi} \int\limits_{0}^{2\pi} \int\limits_{-1}^{1} \mathbf{\Pi}\big(x, \mu, \mu', \varphi - \varphi'\big)\mathbf{I}(x, \mu', \varphi')\mathrm{d}\mu'\mathrm{d}\varphi' + \mathbf{Q}(x, \mu, \varphi).$$

(2.2)

Here, $\mathbf{I}(x, \mu, \varphi)$ is the Stokes vector expressed a function of the polar angle cosine μ (measured from the upward vertical), the azimuth angle φ relative to some fixed direction, and vertical optical thickness (for extinction) x measured from the top of the layer. The 4-vector $\mathbf{I}$ has components $\{I, Q, U, V\}$ where I is the total intensity, Q and U describe linear polarization, and U circular polarization. The degree of polarization is given by $P = I^{-1}\sqrt{Q^2 + U^2 + V^2}$. In addition, $\omega(x)$ is the single scattering albedo and $\mathbf{\Pi}\big(x, \mu, \mu', \varphi - \varphi'\big)$ the phase matrix for scattering of incident Stokes vectors with respect to the meridian plane.

In our formulation, we assume that the atmosphere comprises a stratum of optically uniform layers, so that in any given layer, the optical inputs $\omega(x)$ and $\mathbf{\Pi}\big(x, \mu, \mu', \varphi - \varphi'\big)$ do not depend on the optical thickness x, and we henceforth drop this dependence. If there are N layers in total, with vertical optical depths given by $\{\Delta_k\}, k = 1, \ldots N$, the cumulative optical depth coordinate τ in layer n is related to x through: $\tau = x + \sum_{k=1}^{n} \Delta_k$.

The term $\mathbf{Q}(x, \mu, \varphi)$ is the source function. For solar beam scattering, we have:

$$\mathbf{Q}(x, \mu, \varphi) = \frac{\omega}{4\pi} \mathbf{\Pi}(\mu, -\mu_0, \varphi - \varphi_0)\mathbf{F}_{\odot} T_a \exp[-\lambda x],$$

(2.3)

where $\{-\mu_0, \varphi_0\}$ is the solar direction relative to the meridional plane, $\mathbf{F}_{\odot}$ the solar beam Stokes flux vector before attenuation; in the Earth's atmosphere $\mathbf{F}_{\odot} = \big[F_{\odot}, 0, 0, 0\big]^{\mathrm{T}}$ (natural sunlight, unpolarized). The term $T_a \exp[-\lambda x]$ in Eq. (2.3) is the solar beam attenuation in the pseudo-spherical (P-S) approximation, T_a being the atmospheric transmittance to layer top, and λ a geometrical factor (the "average secant"). The P-S formulation treats solar beam attenuation for a curved atmosphere, but all scattering takes place in a plane-parallel medium; indeed, $\lambda = -1/\mu_0$ when beam attenuation is itself plane-parallel.

In the thermal emission regime (black-body equilibrium), scattering is assumed isotropic and the source term is:

$$\mathbf{Q}(x, \mu, \varphi) = \frac{[1 - \omega]}{4\pi}\mathbf{B}(x).$$

(2.4)

Here, $\mathbf{B}(x) = [B(x), 0, 0, 0]^{\mathrm{T}}$, with the Planck function $B(x)$ expressed as a function of x; in common with other RT models, we will assume a linear dependence $B(x) = a + bx$ across a single layer.

The connection between matrix $\mathbf{\Pi}$ and the scattering matrix $\mathbf{F}(\Theta)$ (assumed to a function only of scattering angle Θ) for vectors defined with respect to the *scattering* plane is given through the application of two 4×4 rotation matrices:

$$\mathbf{\Pi}\big(\mu, \mu', \varphi - \varphi'\big) = \mathbf{L}(\pi - \sigma_2)\mathbf{F}(\Theta)\mathbf{L}(-\sigma_1). \tag{2.5}$$

Rotation angles σ_1 and σ_2 can be expressed in terms of the incident and scatter directions, and the cosine scattering angle is $\cos\Theta = -\mu\mu' + \sqrt{\big(1 - \mu^2\big)}\sqrt{\big(1 - \mu'^2\big)}\cos(\varphi - \varphi')$. For the scalar RTE, matrix $\mathbf{\Pi}$ is replaced by its (1,1) element, the phase function $\Phi(\Theta)$, which is invariant under rotation. We now focus on the scalar LIDORT solutions.

2.2 Scalar RTE Solutions for LIDORT

Considering now the scalar versions of Eqs. (2.1) and (2.2), the first step to solving this RTE is to perform the azimuthal separation, that is, to express the total intensity as a cosine-azimuth series $I(x, \mu, \phi) = \sum_m (2 - \delta_{m0}) I_m(x, \mu) \cos m(\phi - \phi_0)$, for the set of Fourier components $I_m(x, \mu)$. In addition, phase functions $\Phi(\Theta)$ are also separated in azimuth via the addition theorem for Legendre polynomials. [In VLIDORT, the Stokes 4-vector elements are separated in terms of Fourier sine and cosine series in relative azimuth angle, and the azimuthal separation is completed by expressing the elements of $\mathbf{F}(\Theta)$ in terms of generalized spherical functions; see Appendix A, Sect. A.1].

The resulting RTE for component $I_m(x, \mu)$ is then solved by applying a double Gauss-Legendre quadrature approximation of multiple scattering integrals. The quadrature discrete ordinates are $\{\pm\mu_j, c_j\}$, $j = 1, \ldots N_d$, in the positive and negative polar half spaces. This results in the following set of $2N_d$ coupled linear differential equations to be solved in a given layer n:

$$\pm \mu_i \frac{dI_n(x, \pm\mu_i)}{dx} = -I_n(x, \pm\mu_i)$$

$$+ \frac{\omega_n}{2} \sum_{j=1}^{N_d} c_j \big[I_n\big(x, \mu_j\big) \Pi_n\big(\pm\mu_i, \mu_j\big) + I_n\big(x, -\mu_j\big) \Pi_n\big(\pm\mu_i, -\mu_j\big) \big] + Q_n(x, \pm\mu_i).$$

$$\tag{2.6}$$

Here, $\Pi_n\big(\pm\mu, \mu'\big) = \sum_{l=m}^{2N_d-1} \beta_{nl} P_m^l(\pm\mu) P_m^l\big(\mu'\big)$ in terms of associated Legendre functions $P_m^l(\mu)$, and the Fourier index in Eq. (2.6) has been suppressed for ease of exposition. In line with the azimuth separation, the coefficients β_{nl} arise from the phase function Legendre polynomial expansion: $\Phi_n(\Theta) = \sum_{l=0}^{2N_d-1} \beta_{nl} P_l(\cos\Theta)$. The set of properties $\{\Delta_n, \omega_n, \beta_{nl}\}$ are fundamental inputs to the RTE (for the scalar case); these are the IOPs (inherent optical properties), and it is the task of the user to define them for LIDORT (for VLIDORT there will be a set of 6 expansion coefficients—see Appendix A, Sect. A.1).

The RTE Eq. (2.6) is solved by first finding the homogeneous solutions (in the absence of source terms $Q_n(x, \pm\mu_i)$, and then determining the particular integral with the inclusion of the source term. This is a standard procedure, and the result is:

$$I_n(x, \pm\mu_i) = \sum_{\alpha=1}^{N_d} \left[C_{n\alpha}^+ X_{in\alpha}^\pm e^{-k_{n\alpha}x} + C_{n\alpha}^- X_{in\alpha}^\mp e^{-k_{n\alpha}(\Delta_n - x)} \right] + G_{ni}^\pm(x). \qquad (2.7)$$

The homogeneous solutions are found by solving an eigensystem of order N_d; we skip the details here for the scalar case, but Appendix A has the derivation for the vector RTE. Separation constants $\{k_{n\alpha}\}$ and solution vectors $\left\{ \mathbf{X}_{n\alpha}^\pm \right\}$ arise from the eigensystem solution. Quantities $C_{n\alpha}^\pm$ are the constants of integration which are determined through application of the boundary conditions (Sect. 2.3). The presence of the exponential $e^{-k_{n\alpha}(\Delta_n - x)}$ ensures that solutions are numerically bounded (Stamnes and Conklin 1984). The main complication in the vector case is the presence of complex-valued solutions for situations with circular polarization (Appendix A, Sect. A.2).

Solar Sources

Given that the solar source term in Eq. (2.3) has exponential dependence $\exp[-\lambda_n x]$ on optical thickness, it is possible to substitute a trial solution with the same form; eliminating the optical thickness dependence results in a linear-algebra system which can be solved using standard methods. This is the procedure used in VLIDORT (see Appendix A).

In LIDORT however, we use a Green's function method (Siewert 2000a, Spurr 2002), in which the solution is written as follows:

$$G_{ni}^\pm(x) = \sum_{\alpha=1}^{N_d} \left[A_{n\alpha} \mathcal{M}_{n\alpha}^+(x) X_{in\alpha}^\pm + B_{n\alpha} \mathcal{M}_{n\alpha}^-(x) X_{in\alpha}^\mp \right]; \qquad (2.8a)$$

$$A_{n\alpha} = \frac{1}{R_{n\alpha}} \sum_{j=1}^{N_d} c_j \left[Q_{jn}^+ X_{jn\alpha}^+ + Q_{jn}^- X_{jn\alpha}^- \right]; \qquad (2.8b)$$

$$B_{n\alpha} = \frac{1}{R_{n\alpha}} \sum_{j=1}^{N_d} c_j \left[Q_{jn}^- X_{jn\alpha}^+ + Q_{jn}^+ X_{jn\alpha}^- \right]; \qquad (2.8c)$$

$$R_{n\alpha} = \sum_{j=1}^{N_d} \mu_j c_j \left[X_{jn\alpha}^+ X_{jn\alpha}^+ - X_{jn\alpha}^- X_{jn\alpha}^- \right]; \qquad (2.8d)$$

$$\mathcal{M}_{n\alpha}^+(x) = F_\odot T_n \frac{e^{-k_{n\alpha}x} - e^{-\lambda_n x}}{\lambda_n - k_{n\alpha}}; \qquad (2.8e)$$

$$\mathcal{M}_{n\alpha}^-(x) = F_\odot T_n \frac{e^{-\lambda_n x} - e^{-\lambda_n \Delta_n} e^{-k_{n\alpha}(\Delta_n - x)}}{\lambda_n + k_{n\alpha}}; \qquad (2.8f)$$

$$Q_{jn}^{\pm} = (2 - \delta_{m0})\frac{\omega_n}{2} \sum_{l=m}^{LM} P_m^l(\pm\mu_j)\beta_{ln} P_m^l(-\mu_0). \qquad (2.8g)$$

Terms $\left\{A_{n\alpha}, B_{n\alpha}, R_{n\alpha}, Q_{jn}^{\pm}\right\}$ are independent of optical thickness x in layer n, depending only on the homogeneous solution vectors $\mathbf{X}_{n\alpha}^{\pm}$, the discrete ordinate quadrature and the IOPs $\{\omega_n, \beta_{ln}\}$ for a given layer n. The optical thickness dependency is expressed through _multipliers_ $\mathcal{M}_{n\alpha}^{\pm}(x)$, where $T_n \exp[-\lambda_n x]$ is the solar beam attenuation in the (P-S) approximation.

The Green's function treatment contains only closed-form algebraic expressions. The solution is also stable: in the limit $\lambda_n \to k_{n\alpha}$, $\mathcal{M}_{n\alpha}^{+}(x) \to F_{\odot}T_n\lambda_n x$ in Eq. (2.8e), though a Taylor series expansion may be required close to this limit.

Thermal Sources

With a source term given by Eq. (2.4), and isotropic scattering, only the Fourier $m = 0$ term survives for thermal sources. We take the linear dependence of the Planck function on x:

$$B_n(x) = a_n + b_n x; \quad a_n = B_{n-1}; \quad b_n = \frac{B_n - B_{n-1}}{\Delta_n}; \qquad (2.9)$$

in terms of Planck functions B_{n-1} and B_n at layer boundaries, and the set $\{B_n\}$ for $n = 0, 1, \ldots N$ is a fundamental thermal-emission input to the RT models. The surface emission term B_{Surf} is also an input.

It is possible to substitute a trial solution of the same linear form as a function of x. This is the procedure adopted for VLIDORT; removal of x-dependency again leads to linear algebra solutions for the particular integral.

In LIDORT however, we use the closed form Green's function solution, which has the following form (Spurr 2002):

$$\mathbf{G}_n^{\pm}(x) = \varpi_n \sum_{\alpha=1}^{N_d} \left[A_{n\alpha}^{-}\mathcal{T}_{n\alpha}^{-}(x)\mathbf{X}_{n\alpha}^{\mp} + A_{n\alpha}^{+}\mathcal{T}_{n\alpha}^{+}(x)\mathbf{X}_{n\alpha}^{\pm}\right];$$

$$A_{n\alpha}^{\pm} = \frac{1}{R_{n\alpha}} \sum_{j=1}^{N_d} c_j \left[X_{jn\alpha}^{+} \pm X_{jn\alpha}^{-}\right]. \qquad (2.10)$$

Here, $R_{n\alpha}$ is defined in Eq. (2.8d), and $\varpi_n = 1 - \omega_n$.

The multipliers $\mathcal{T}_{n\alpha}^{\pm}(x)$ are easy to evaluate, and for reasons that will become apparent when we consider Planck function linearization in Sect. 3.3, we note only the values at $x = \Delta_n$ and $x = 0$ respectively, expressed as linear combinations of B_{n-1} and B_n:

$$\mathcal{T}_{n\alpha}^{+}(\Delta_n) = \mathcal{P}_{n\alpha}^{+} B_{n-1} + \mathcal{Q}_{n\alpha}^{+} B_n; \quad \mathcal{T}_{n\alpha}^{-}(0) = \mathcal{P}_{n\alpha}^{-} B_{n-1} + \mathcal{Q}_{n\alpha}^{-} B_n. \qquad (2.11a)$$

$$\mathcal{P}_{n\alpha}^{+} = \mathcal{S}_{n\alpha} - \frac{1}{\Delta_n k_{n\alpha}}(\Delta_n - \mathcal{S}_{n\alpha}); \quad \mathcal{Q}_{n\alpha}^{+} = +\frac{1}{\Delta_n k_{n\alpha}}(\Delta_n - \mathcal{S}_{n\alpha}); \quad (2.11b)$$

$$\mathcal{P}_{n\alpha}^{-} = \mathcal{S}_{n\alpha} + \frac{1}{\Delta_n k_{n\alpha}}\big(\Delta_n e^{-\Delta_n k_{n\alpha}} - \mathcal{S}_{n\alpha}\big); \quad \mathcal{Q}_{n\alpha}^{-} = -\frac{1}{\Delta_n k_{n\alpha}}\big(\Delta_n e^{-\Delta_n k_{n\alpha}} - \mathcal{S}_{n\alpha}\big).$$

$$(2.11c)$$

In these equations, $\mathcal{S}_{n\alpha} = \frac{1}{k_{n\alpha}}\big(1 - e^{-\Delta_n k_{n\alpha}}\big)$. We note also that $\mathcal{T}_{n\alpha}^{+}(0) = 0$; $\mathcal{T}_{n\alpha}^{-}(\Delta_n) = 0$.

We can now compute the particular integrals at layer boundaries, by substituting these results into Eq. (2.7); again we write the answers as linear combinations of Planck functions:

$$\mathbf{G}_n^{\pm}(0) = B_{n-1}\mathbf{P}_{n\alpha}^{\pm} + B_n\mathbf{Q}_{n\alpha}^{\pm}$$

$$\equiv B_{n-1}\varpi_n \sum_{\alpha=1}^{N}\big[A_{n\alpha}^{-}\mathcal{P}_{n\alpha}^{-}\mathbf{X}_{n\alpha}^{\mp}\big] + B_n\varpi_n \sum_{\alpha=1}^{N}\big[A_{n\alpha}^{-}\mathcal{Q}_{n\alpha}^{-}\mathbf{X}_{n\alpha}^{\mp}\big]; \quad (2.12a)$$

$$\mathbf{G}_n^{\pm}(\Delta_n) = B_{n-1}\mathbf{R}_{n\alpha}^{\pm} + B_n\mathbf{S}_{n\alpha}^{\pm}$$

$$\equiv B_{n-1}\varpi_n \sum_{\alpha=1}^{N}\big[A_{n\alpha}^{+}\mathcal{P}_{n\alpha}^{+}\mathbf{X}_{n\alpha}^{\pm}\big] + B_n\varpi_n \sum_{\alpha=1}^{N}\big[A_{n\alpha}^{+}\mathcal{Q}_{n\alpha}^{+}\mathbf{X}_{n\alpha}^{\pm}\big]. \quad (2.12b)$$

Although we have chosen values at the layer boundaries, the multipliers $\mathcal{T}_{n\alpha}^{\pm}(x)$ can be found readily enough for arbitrary optical thickness x, and quantities $\mathbf{G}_n^{\pm}(x)$ can then be derived as linear combinations of Planck functions as in Eq. (2.12).

2.3 Boundary Conditions and Post-processed Solutions

Confining our attention to the scalar case, we summarize the main steps to complete intensity field computations. The constants of integration $\{C_{n\alpha}^{\pm}\}$ in Eq. (2.7) are determined by imposition of three boundary conditions. These are: (I) for $n = 1$, there is no downwelling diffuse field at the top of the atmosphere (TOA); (II) upwelling and downwelling intensity fields are continuous across all intermediate layer boundaries; (III) there is a surface reflection condition at the bottom of the atmosphere (BOA) linking the upwelling and downwelling intensity fields there. Explicitly for the scalar case with solar beam sources:

$$\sum_{\alpha=1}^{N_d}\big[C_{n\alpha}^{+}\mathbf{X}_{n\alpha}^{+} + C_{n\alpha}^{-}\mathbf{X}_{n\alpha}^{-}\Gamma_{n\alpha}\big] = -\mathbf{G}_n^{\pm}(0); \quad (n = 1) \quad (2.13a)$$

$$\sum_{\alpha=1}^{N_d} \left[C_{p\alpha}^+ \mathbf{X}_{p\alpha}^\pm \Gamma_{p\alpha} + C_{p\alpha}^- \mathbf{X}_{p\alpha}^\mp \right] - \sum_{\alpha=1}^{N_d} \left[C_{n\alpha}^+ \mathbf{X}_{n\alpha}^\pm + C_{n\alpha}^- \mathbf{X}_{n\alpha}^\mp \Gamma_{n\alpha} \right]$$
$$= \sum_{q=1}^{N_Q} \left[\mathbf{G}_n^\pm(0) - \mathbf{G}_p^\pm(\Delta_p) \right]; \qquad (2.13b)$$

$$\sum_{\alpha=1}^{N_d} \left[C_{N\alpha}^+ \mathbf{U}_\alpha^- \Gamma_{N\alpha} + C_{N\alpha}^- \mathbf{U}_\alpha^+ \right] = -\mathbf{H} + \mathbf{H}_\odot; \qquad (2.13c)$$

$$\mathbf{U}_\alpha^\pm = \mathbf{X}_{n\alpha}^\pm - (2 - \delta_{m0}) \mathbf{R} \mathbf{L} \mathbf{X}_{n\alpha}^\mp; \qquad (2.13d)$$

$$\mathbf{H} = \mathbf{G}_N^-(\Delta_N) - (2 - \Delta_{m0}) \mathbf{R} \mathbf{L} \mathbf{G}_N^+(\Delta_N);$$

$$\mathbf{H}_\odot = (2 - \Delta_{m0}) \left(\frac{F_\odot \mu_0}{4\pi} \right) T_\odot \mathbf{R}_\odot \qquad (2.13e)$$

Here, $\Gamma_{n\alpha} \equiv e^{-k_{n\alpha}\Delta_n}$, and in Eq. (2.13b) for the intermediate boundaries, $p = n-1$ and $n = 2, \ldots N$, where there are N layers in all. In the surface condition (2.13c), we have some auxiliary definitions in (2.13d) and (2.13e). In these equations, $\mathbf{L}$ is a diagonal matrix of rank N_d with entries $\{\mu_i c_i\}, i = 1, \ldots N_d$, $\mathbf{R}$ is a surface reflection matrix of rank N_d, and $\mathbf{H}_\odot$ is the direct beam reflectance for solar flux $F_\odot$, whole-atmosphere beam transmittance $T_\odot$, and beam surface reflectance $\mathbf{R}_\odot$. Reflection matrices for a BRDF surface are constructed from knowledge of the surface bidirectional reflectances, which are provided by the BRDF supplements (Sect. 4.2). For a Lambertian surface, $\mathbf{R} = \mathbf{R}_\odot = A_L \mathbf{E}$, in terms of the Lambertian albedo A_L ($\mathbf{E}$ is the identity matrix of rank N_d), and only the Fourier $m = 0$ term survives on the right of Eq. (2.13c).

The resulting *boundary value problem* (BVP) is well known from the literature. Combining these conditions, the BVP is a sparse banded tri-diagonal linear algebra system:

$$\mathbb{A}\mathbb{C} = \mathbb{B}; \;\rightarrow \mathbb{C} = \mathbb{A}^{-1}\mathbb{B}. \qquad (2.14)$$

Here, $\mathbb{C}$ is a vector with elements $\left\{ C_{n\alpha}^+, C_{n\alpha}^- \right\}, \alpha = 1, \ldots N_d, n = 1, \ldots N$, $\mathbb{A}$ is a matrix constructed from quantities $\left\{ \mathbf{X}_{n\alpha}^\pm, \mathbf{U}_\alpha^\pm, \Gamma_{n\alpha} \right\}$, and $\mathbb{B}$ is the column vector constructed from solutions on the right hand sides of Eqs. (2.13a). A visualization of this algebra system is given in the LIDORT work (Spurr et al. 2001). The solution for $\mathbb{C}$ is obtained using a band compression algorithm and linear algebra modules from the LAPACK numerical package. The procedure for the vector case is similar (Spurr 2006).

The BVP yields the discrete ordinate fields $I_{ni}^\pm(x), i = 1, \ldots N_d$ everywhere. These fields are sufficient to determine integrated quantities such as the actinic flux. To obtain output at other polar angles, we apply source function integration, in which the discrete ordinate field is substituted in all multiple-scatter integrals in the original RTE, and the "post-processed" field is then developed recursively by optical depth integration starting from TOA (downwelling radiation) or BOA (upwelling). Here,

we write down the main results based on the Green's function method, looking just at the upwelling intensity $I_n^-(x, \mu)$ for layer n in polar direction μ, with $I_{n-1}^-(\mu)$ denoting the value of the field at layer upper boundary ($x = 0$).

$$I_{n-1}^-(\mu) = I_n^-(\mu) \exp\left[-\frac{\Delta_n}{\mu}\right] + S_n^-(\mu);$$

$$S_n^-(\mu) = H_n^-(\mu) + E_n^-(\mu) + D_n^-(\mu). \tag{2.15}$$

$$H_n^-(\mu) = \sum_{\alpha=1}^{N_d} \left[C_{n\alpha}^+ Z_{n\alpha}^+(\mu)\mathcal{H}_{n\alpha}^{--}(\mu) + C_{n\alpha}^- Z_{n\alpha}^-(\mu)\mathcal{H}_{n\alpha}^{-+}(\mu)\right]; \tag{2.16a}$$

$$Z_{n\alpha}^\pm(\mu) = \frac{1}{2}\sum_{l=m}^{2N_d-1} \omega_n \beta_{nl} P_m^l(\mu) \sum_{j=1}^{N_d} P_m^l(\mp\mu_j) c_j X_{jn\alpha}^\pm; \tag{2.16b}$$

$$\mathcal{H}_{n\alpha}^{-+}(\mu) = \frac{e^{-k_{n\alpha}\Delta_n} - e^{-\Delta_n/\mu}}{1 - \mu k_{n\alpha}}; \quad \mathcal{H}_{n\alpha}^{--}(\mu) = \frac{1 - e^{-k_{n\alpha}\Delta_n}e^{-\Delta_n/\mu}}{1 + \mu k_{n\alpha}}; \tag{2.16c}$$

$$E_n^-(\mu) = Q_n^-(\mu)\mathcal{E}_n^-(\mu); \quad \mathcal{E}_n^-(\mu) = \frac{1 - e^{-\lambda_n\Delta_n}e^{-\Delta_n/\mu}}{1 + \mu\lambda_n}; \tag{2.16d}$$

$$Q_n^\pm(\mu) = \frac{1}{2}\sum_{l=m}^{2N_d-1} \omega_n \beta_{nl} P_m^l(\mu) \sum_{j=1}^{N_d} P_m^l(\mp\mu_j) c_j Q_{jn}^\pm; \tag{2.16e}$$

$$D_n^-(\mu) = \sum_{\alpha=1}^{N_d} \left[Z_{n\alpha}^+(\mu) A_{n\alpha} \mathcal{D}_{n\alpha}^{-+}(\mu) + Z_{n\alpha}^-(\mu) B_{n\alpha} \mathcal{D}_{n\alpha}^{--}(\mu)\right]; \tag{2.16f}$$

$$\mathcal{D}_{n\alpha}^{-+}(\mu) = \frac{\mathcal{H}_{n\alpha}^{-+}(\mu) - \mathcal{E}_n^-(\mu)}{\lambda_n - k_{n\alpha}}; \quad \mathcal{D}_{n\alpha}^{--}(\mu) = \frac{\mathcal{E}_n^-(\mu) - \mathcal{H}_{n\alpha}^{--}(\mu)e^{-\lambda_n\Delta_n}}{\lambda_n + k_{n\alpha}}. \tag{2.16g}$$

Similar results apply for the downwelling recursion. The quantities $Z_{n\alpha}^\pm(\mu)$ can be thought of as solution vectors at polar stream μ. The groups $\left\{\mathcal{H}_{n\alpha}^{-\pm}(\mu), \mathcal{E}_n^-(\mu), \mathcal{D}_{n\alpha}^{-\pm}(\mu)\right\}$ of post-processing multipliers all arise from integrations over optical depth across the layer n. For more discussion, see (Spurr 2002). With the multiplier $\mathcal{D}_{n\alpha}^{-+}(\mu)$ in Eq. (2.16g), both numerator and denominator tend to zero in the limit $\lambda_n \rightarrow k_{n\alpha}$, and this kind of instability has to be treated with Taylor series expansions [c.f. Eq. (2.8e)]. Such expansions are treated in detail in Appendix C.

Thermal Sources. The boundary-value problem for thermal sources is constructed in a similar fashion. For simplicity, we assume a dark surface with no reflection; surface emissivity is then unity (in the following, B_{surf} is the surface emission Planck function). Writing the BVP as $\mathbb{A}\mathbb{C} = \mathbb{B}$ once again, the explicit form for $\mathbb{B}$ may be written:

$$\mathbb{B} = \begin{pmatrix} -\mathbf{G}_1^+(0) \\ \mathbf{G}_2^{\pm}(0) - \mathbf{G}_1^{\pm}(\Delta_1) \\ \dots \\ \mathbf{G}_N^{\pm}(0) - \mathbf{G}_{N-1}^{\pm}(\Delta_{N-1}) \\ \mathbf{G}_N^-(\Delta_N) + \mathbf{E}B_{surf} \end{pmatrix}$$

$$= \begin{pmatrix} -B_0\mathbf{P}_{1\alpha}^+ - B_1\mathbf{Q}_{1\alpha}^+ \\ -B_2\mathbf{Q}_{2\alpha}^{\pm} - B_0\mathbf{R}_{1\alpha}^{\pm} + B_1\left(\mathbf{P}_{2\alpha}^{\pm} - \mathbf{S}_{1\alpha}^{\pm}\right) \\ \dots \\ -B_N\mathbf{Q}_{N\alpha}^{\pm} - B_{N-2}\mathbf{R}_{N-1,\alpha}^{\pm} + B_{N-1}\left(\mathbf{P}_{N\alpha}^{\pm} - \mathbf{S}_{N-1,\alpha}^{\pm}\right) \\ B_{N-1}\mathbf{R}_{N\alpha}^- + B_N\mathbf{S}_{N\alpha}^- + \mathbf{E}B_{surf} \end{pmatrix}. \tag{2.17}$$

The notation follows definitions used in Eq. (2.12). Here, $\mathbf{E}$ is vector of rank N_d with unit entries; this represents the emissivity. We note that each of these quantities is dependent on the complete set $\{B_n\}, n = 0, \dots N$, and also on B_{surf}. We omit derivation of the post-processed thermal solutions, only remarking that they may be readily expressed as linear combinations of atmospheric and surface Planck function inputs.

2.4 Linearization of the Intensity Field

Continuing with the scalar LIDORT model, we will mostly quote the linearization results from Spurr (2002). The treatment here is for profile linearizations, where we recognize that some particular integrals in any given layer n will depend upon atmospheric parameters not only in that layer, but also in other layers $p \neq n$; in other words, the linearization will involve "cross-layer derivatives". These cross-layer derivatives arise only from solar beam attenuation through layers above and including n, that is, for $p \leq n$. In general, we the "dot" notation:

$$\dot{A}_{pn\alpha} \equiv \xi_p \frac{\partial A_{n\alpha}}{\partial \xi_p} \tag{2.18}$$

for the *normalized* partial derivative of quantity $A_{n\alpha}$ in layer n with respect to atmospheric variable ξ_p in layer p.

For homogeneous solutions, there are no cross-layer derivatives. In this case, it is necessary to differentiate the eigenproblem that defines solutions to the homogeneous equations. [For VLIDORT, linearization of the vector homogeneous RTE solutions is trickier, thanks to the presence of complex eigensolutions—this linearization is

found in Appendix A, along with analytic differentiation of the solar-beam particular integral obtained by substitution methods].

Continuing for LIDORT, we assume that the following linearizations already exist:

$$\mathbf{Y}_{n\alpha}^{\pm} = \dot{\mathbf{X}}_{n\alpha}^{\pm} \equiv \xi_p \frac{\partial \mathbf{X}_{n\alpha}^{\pm}}{\partial \xi_p}; \quad f_{n\alpha} = \dot{k}_{n\alpha} \equiv \xi_p \frac{\partial k_{n\alpha}}{\partial \xi_p}. \tag{2.19}$$

Now consider the Green's function solution driven by source terms dependent on quantities $\{Q_{ni}^{\pm}, \lambda_n\}$; again we define the linearizations: $\dot{Q}_{pni}^{\pm} = \xi_p \frac{\partial Q_{ni}^{\pm}}{\partial \xi_p}; \dot{\lambda}_{pn} = \xi_p \frac{\partial \lambda_n}{\partial \xi_p}$, and suppose that they have been established from differentiation of the appropriate source terms; details may be found in Spurr (2002).

First, we differentiate analytically some of the results in Eq. (2.8a) to find:

$$\dot{A}_{pn\alpha} = \frac{\left(\sum_{j=1}^{N_d} c_j \left[Q_{nj}^+ Y_{jn\alpha}^+ + Q_{nj}^- Y_{jn\alpha}^- \right] - \dot{R}_{n\alpha} A_{n\alpha} \right) \delta_{np}}{R_{n\alpha}} + \frac{\sum_{j=1}^{N_d} \mu_j c_j \left[X_{jn\alpha}^+ \dot{Q}_{pni}^+ + X_{jn\alpha}^- \dot{Q}_{pni}^- \right]}{R_{n\alpha}};$$

$$\dot{R}_{n\alpha} = 2 \sum_{j=1}^{N_d} \mu_j c_j \left[X_{jn\alpha}^+ Y_{jn\alpha}^+ + X_{jn\alpha}^- Y_{jn\alpha}^- \right]. \tag{2.20}$$

A similar expression pertains for $\dot{B}_{pn\alpha}$. Note that the first term in $\dot{A}_{pn\alpha}$ is only present for $p = n$; the second term has additional cross-derivatives.

Analytic differentiation of the first of the multiplier terms in Eq. (2.8e) yields:

$$\dot{\mathcal{M}}_{pn\alpha}^+(x) = -\mathcal{M}_{n\alpha}^+(x) \left[\frac{\dot{T}_{pn}}{T_n} - \frac{\left(\dot{\lambda}_{pn} - f_{n\alpha} \delta_{np} \right)}{\lambda_n - k_{n\alpha}} \right]$$

$$+ \frac{-(\dot{x} k_{n\alpha} + x f_{n\alpha}) e^{-k_{n\alpha} x} \delta_{np} - \left(\dot{x} \lambda_n \delta_{np} + x \dot{\lambda}_{pn} \right) e^{-\lambda_n x}}{\lambda_n - k_{n\alpha}}. \tag{2.21}$$

A similar exercise pertains for the other multiplier derivative $\dot{\mathcal{M}}_{pn\alpha}^-(x)$ (not done here). Note that since x is a partial optical thickness in layer n, we have $\dot{x} = x(\dot{\Delta}_n / \Delta_n)$, where $\dot{\Delta}_n \equiv \xi_n \frac{\partial \Delta_n}{\partial \xi_n}$ is one of the fundamental linearized optical property inputs. Note that the linearized multiplier $\dot{\mathcal{M}}_{pn\alpha}^+(x)$ in Eq. (2.21) will be subject to instability as $\lambda_n \to k_{n\alpha}$, and we will again require Taylor series expansions in this limit (see Appendix C).

Continuing, the final linearized Green's function solution for one of the source term series is obtained by chain-rule differentiation of Eq. (2.8a), using results obtained above:

$$\dot{\mathbf{G}}_{pn}^{\pm}(x) = \sum_{\alpha=1}^{N_d} \left[\dot{A}_{pn\alpha} \mathcal{M}_{n\alpha}^+(x) \mathbf{X}_{n\alpha}^{\pm} + A_{n\alpha} \dot{\mathcal{M}}_{pn\alpha}^+(x) \mathbf{X}_{n\alpha}^{\pm} + \delta_{np} A_{n\alpha} \mathcal{M}_{n\alpha}^+(x) \mathbf{Y}_{n\alpha}^{\pm} \right]$$

$$+ \sum_{\alpha=1}^{N_d} \left[\dot{B}_{pn\alpha} \mathcal{M}_{n\alpha}^-(x) \mathbf{X}_{n\alpha}^{\mp} + B_{n\alpha} \dot{\mathcal{M}}_{pn\alpha}^-(x) \mathbf{X}_{n\alpha}^{\mp} + \delta_{np} B_{n\alpha} \mathcal{M}_{n\alpha}^-(x) \mathbf{Y}_{n\alpha}^{\mp} \right]. \tag{2.22}$$

Next, we can through-differentiate Eq. (2.7) to obtain Jacobians for the discrete ordinate radiation field, once we have obtained derivatives of the integration constants. Formal linearization of Eq. (2.14) yields (keeping our abbreviated "dot" derivative notation to represent the derivative operator $\xi_p \frac{\partial}{\partial \xi_p}$):

$$\dot{\mathbb{C}} = \mathbb{A}^{-1}\mathbb{B}'; \quad \mathbb{B}' = \dot{\mathbb{B}} - \dot{\mathbb{A}}\mathbb{C}. \tag{2.23}$$

Thus, the linearized BVP is completed through back-substitution with a different column vector $\mathbb{B}'$; we already have the inverse $\mathbb{A}^{-1}$ from solving the original BVP. For the linearization of $\mathbb{A}$, we recall that this matrix is made up from elements $\left\{\mathbf{X}_{n\alpha}^{\pm}, k_{n\alpha}\right\}$, for which we already have derivatives [Eq. (2.19)], so we assume this has been done; this part of the linearization is the same regardless of source terms, and more details on deriving $\dot{\mathbb{A}}$ in Eq. (2.23) may be found in the LIDORT papers.

For the linearized vector $\dot{\mathbb{B}}$ in Eq. (2.23), we observe that $\mathbb{B}$ is constructed from vectors $\left\{\mathbf{G}_n^{\pm}(x)\right\}$ in all layers, and these linearizations are given above in Eqs. (2.20) and (2.21); values of x are required only at layer boundaries ($x = 0$ or $x = \Delta_n$). The only new consideration comes with the direct beam term in Eq. (2.13c), which has the linearization [see Eq. (2.13e)]:

$$\dot{\mathbf{H}}_{\odot} = (2 - \delta_{m0})\left(\frac{F_{\odot}\mu_0}{4\pi}\right)\dot{T}_{\odot}\mathbf{R}_{\odot}. \tag{2.24}$$

In a plane-parallel atmosphere, $T_{\odot} = \exp\left[-\frac{1}{\mu_0}\sum_{n=1}^{N}\Delta_n\right]$, so that $\dot{T}_{\odot} = -\frac{T_{\odot}}{\mu_0}\dot{\Delta}_p$. The linearization of $T_{\odot}$ in the pseudo-spherical approximation is a little more complicated (Sect. 3.4). The linearized BVP solution gives derivatives of the integration constants, that is, $\dot{C}_{pn\alpha}^{\pm} = \xi_p \frac{\partial C_{n\alpha}^{\pm}}{\partial \xi_p}$.

For linearization of the post-processed solution, we must differentiate Eq. (2.15):

$$\dot{I}_{p,n-1}^{-}(\mu) = \left[\dot{I}_{pn}^{-}(\mu) - \delta_{np}\frac{\dot{\Delta}_n}{\mu}I_n^{-}(\mu)\right]\exp\left(-\frac{\Delta_n}{\mu}\right) + \dot{S}_{pn}^{-}(\mu), \tag{2.25}$$

and the system in Eqs. (2.16); this is a lengthy but straightforward exercise, and here we indicate a few of the key steps in this process. The three contributions to the integrated source terms $\dot{S}_{pn}^{-}(\mu)$ may all be differentiated using the chain rule and quantities that have already been linearized; we note just the following linearizations (the rest can be inferred).

$$\dot{\mathcal{H}}_{n\alpha}^{+}(\mu) = \delta_{np}\frac{\left[-e^{-k_{n\alpha}\Delta_n}\left(f_{n\alpha}\Delta_n + k_{n\alpha}\dot{\Delta}_n\right) + \frac{\dot{\Delta}_n}{\mu}e^{-\Delta_n/\mu}\right] + \mu f_{n\alpha}\mathcal{H}_{n\alpha}^{+}(\mu)}{1 - \mu k_{n\alpha}}; \tag{2.26a}$$

$$\dot{Z}_{n\alpha}^{\pm}(\mu) = \frac{\delta_{np}}{2} \sum_{l=m}^{2N_d-1} \omega_n \beta_{nl} P_m^l(\mu) \sum_{j=1}^{N_d} P_m^l(\mp \mu_j) c_j Y_{jn\alpha}^{\pm}$$

$$+ \frac{\delta_{np}}{2} \sum_{l=m}^{2N_d-1} \left[\dot{\omega}_n \beta_{nl} + \omega_n \dot{\beta}_{nl} \right] P_m^l(\mu) \sum_{j=1}^{N_d} P_m^l(\mp \mu_j) c_j X_{jn\alpha}^{\pm}; \qquad (2.26b)$$

$$\dot{E}_{np}^{-}(\mu) = \left[\dot{Q}_{pn}^{-}(\mu) \mathcal{E}_n^{-}(\mu) + Q_n^{-}(\mu) \dot{\mathcal{E}}_{pn}^{-}(\mu) \right]; \qquad (2.26c)$$

$$\dot{\mathcal{E}}_{pn}^{-}(\mu) = \frac{\left[\dot{\lambda}_{pn} \Delta_n + \delta_{np}(\lambda_n + 1/\mu) \dot{\Delta}_n \right] e^{-\lambda_n \Delta_n} e^{-\Delta_n/\mu} - \mu \dot{\lambda}_{pn} \mathcal{E}_{nq}^{-}(\mu)}{1 + \mu \lambda_n}; \qquad (2.26d)$$

$$\dot{\mathcal{D}}_{pn\alpha}^{-+}(\mu) = \frac{\delta_{np} \dot{\mathcal{H}}_{n\alpha}^{+}(\mu) - \dot{\mathcal{E}}_{pn}^{-}(\mu) - \left(\dot{\lambda}_{pn} - \delta_{np} f_{n\alpha} \right) \mathcal{D}_{n\alpha}^{-+}(\mu)}{\lambda_n - k_{n\alpha}}. \qquad (2.26e)$$

Other derivatives $\left\{ \dot{\mathcal{H}}_{n\alpha}^{-}(\mu), \dot{\mathcal{D}}_{pn\alpha}^{--}(\mu) \right\}$ may be calculated similarly. Taylor series expansions will be required in the limit $\lambda_n \to k_{n\alpha}$ in Eq. (2.26e). Note that $\dot{\mathcal{H}}_{n\alpha}^{\pm}(\mu)$ and $\dot{Z}_{n\alpha}^{\pm}(\mu)$ have no cross-derivative contributions. Derivatives $\dot{A}_{pn\alpha}$ and $\dot{B}_{pn\alpha}$ have been determined as part of the Green's function linearization, and the derivatives $\dot{Q}_{pn}^{\pm}(\mu)$ emerge from differentiation of Eq. (2.16e). It is clear from this analysis that *linearized IOPs* $\left\{ \dot{\Delta}_n, \dot{\omega}_n, \dot{\beta}_{nl} \right\}$ are also fundamental inputs to the Jacobian calculations; Sect. 3.1 has some examples for their preparation.

This completes the Green's function solution and its linearization for profile Jacobians. Treatment of whole-atmosphere ("column") Jacobians is very similar, and the main equations are the same, with the chief difference being the presence of cross-layer derivatives in all layers. For profile Jacobians, not all layer source terms in Eqs. (2.8) and (2.16) have cross-derivatives; for the column linearization, every equivalent term will have cross-derivatives.

We now turn to surface property linearization. Here, we use the "double-dot" operator notation $\ddot{A} \equiv \frac{\partial A}{\partial \sigma}$ to denote partial differentiation with respect to some surface property σ (which might be the Lambertian albedo, or a parameter that characterizes a BRDF surface, e.g. wind speed for ocean glitter BRDF). Un-normalized derivatives are used for this analysis.

First, we note that eigenproblem variables $\left\{ k_{n\alpha}, \mathbf{X}_{n\alpha}^{\pm} \right\}$ for RTE homogeneous solutions have no surface derivatives $\left(\ddot{k}_{n\alpha} = 0 \right)$; the same applies to the Green's function solutions in Eqs. (2.8) and (2.16). The linearized BVP for the surface case is then

$$\ddot{\mathbb{C}} = \mathbb{A}^{-1} \mathbb{B}''; \quad \mathbb{B}'' = \ddot{\mathbb{B}} - \ddot{\mathbb{A}} \mathbb{C}. \qquad (2.27)$$

For matrix $\ddot{\mathbb{A}}$ and vector $\ddot{\mathbb{B}}$, surface derivatives will only appear for the reflection condition in Eq. (2.13c). Specifically, for the quantities in Eqs. (2.13d) and (2.13e), we have:

$$\ddot{\mathbf{U}}_{\alpha}^{\pm} = -(2 - \delta_{m0}) \mathbf{R} \ddot{\mathbf{L}} \mathbf{X}_{n\alpha}^{\mp}; \qquad (2.28a)$$

$$\ddot{\mathbf{H}} = \ddot{\mathbf{G}}_N^-(\Delta_N) - (2 - \delta_{m0})\left\{\ddot{\mathbf{R}}\mathbf{L}\mathbf{G}_N^+(\Delta_N) + \mathbf{R}\mathbf{L}\ddot{\mathbf{G}}_N^+(\Delta_N)\right\}; \qquad (2.28b)$$

$$\ddot{\mathbf{H}}_\odot = (2 - \delta_{m0})\left(\frac{\mathrm{F}_\odot\mu_0}{4\pi}\right)\mathrm{T}_\odot\ddot{\mathbf{R}}_\odot. \qquad (2.28c)$$

The surface property linearizations $\ddot{\mathbf{R}}$ and $\ddot{\mathbf{R}}_\odot$ are crucial inputs to LIDORT. For the Lambertian case, $\mathbf{R} = \mathbf{R}_\odot = A_L\mathbf{E}$, the surface parameter $\sigma = A_L$, and so $\ddot{\mathbf{R}} = \ddot{\mathbf{R}}_\odot = \mathbf{E}$. For a BRDF surface, it is the task of the BRDF supplement to provide $\ddot{\mathbf{R}}$ and $\ddot{\mathbf{R}}_\odot$ as inputs to the model. Once solved, the linearized BVP gives derivatives $\ddot{C}_{n\alpha}^\pm = \frac{\partial C_{n\alpha}^\pm}{\partial\sigma}$ of the integration constants.

For the post processed solution, we proceed as before, differentiating first Eq. (2.15):

$$\ddot{I}_{n-1}^-(\mu) = \ddot{I}_n^-(\mu)\exp\left(-\frac{\Delta_n}{\mu}\right) + \ddot{S}_n^-(\mu). \qquad (2.29)$$

For $\ddot{S}_n^-(\mu)$, only the term $H_n^-(\mu)$ in Eq. (2.16a) will have surface-property derivatives, thanks to the presence of linearized integration constants $\ddot{C}_{n\alpha}^\pm$; all other quantities have no dependence on surface parameters. This completes the surface linearization process.

3 Aspects of the LIDORT Models

3.1 Preparation of Optical Inputs

As noted already, the LIDORT and VLIDORT models require IOP ("inherent optical property") inputs $\{\Delta_n, \omega_n, \beta_{nl}\}$ and $\{\Delta_n, \omega_n, \mathbf{\Psi}_{nl}\}$ respectively, defined for each layer n. Here, Δ_n is the layer optical thickness for extinction, and ω_n and β_{nl} are the layer single-scattering albedo and the set of phase function Legendre expansion coefficients respectively; $\{\mathbf{\Psi}_{nl}\}$ is the matrix of "Greek coefficients" needed for VLIDORT's generalized spherical function expansions.

Without loss of generality, we confine ourselves to a typical example for the scalar case. In terms of the layer trace-gas absorption optical depth G_n, the Rayleigh scattering optical depth R_n and the aerosol extinction and scattering optical depths ε_n and s_n, we have:

$$\Delta_n = G_n + R_n + \epsilon_n; \quad \omega_n = \frac{R_n + s_n}{\Delta_n}; \quad \beta_{nl} = \frac{R_n\beta_l^{(Ray)} + s_n\beta_{nl}^{(Aer)}}{R_n + s_n}. \qquad (3.1)$$

Here, $\beta_l^{(Ray)}$ and $\beta_{nl}^{(Aer)}$ are the phase function expansion coefficients for Rayleigh and aerosol scattering respectively. For Rayleigh scattering, $\beta_0^{(Ray)} = 1$, and $\beta_2^{(Ray)} = (1 - \rho_r)/(2 + \rho_r)$ in terms of the depolarization ratio ρ_r, with all other coefficients

Table 2 Greek coefficients for Rayleigh scattering

Coefficient	α_l	β_l	γ_l	δ_l	ϵ_l	ζ_l
$l = 0$	0	1	0	0	0	0
$l = 1$	0	0	0	$\frac{3(1-2\rho_r)}{(2+\rho_r)}$	0	0
$l = 2$	$\frac{6(1-\rho_r)}{(2+\rho_r)}$	$\frac{(1-\rho_r)}{(2+\rho_r)}$	$-\frac{\sqrt{6}(1-\rho_r)}{(2+\rho_r)}$	0	0	0

are zero. For aerosol scattering, $\beta_{nl}^{(Aer)}$ may be returned from a Mie or T-matrix (or other) electromagnetic scattering calculation; a commonly-used coefficient set is $\beta_{nl}^{(HG)} = (2l + 1)(g_n)^l$ for the Henyey-Greenstein phase function, where g_n is the layer asymmetry factor.

We suppose the trace gas absorption optical depth for a single absorber is $G_n = \chi_n \alpha_n$, where χ_n is the layer column density, and α_n the absorber cross-section. The Rayleigh scattering optical depth is $R_n = a_n \sigma_{Ray}$ in terms of the layer air column density a_n and Rayleigh cross-section σ_{Ray}. In Eq. (3.1), ϵ_n and s_n are the aerosol layer extinction and scattering optical depths respectively.

For VLIDORT, coefficients $\beta_l^{(Ray)}$ and $\beta_{nl}^{(Aer)}$ in Eq. (3.1) are replaced by the respective Greek matrices $\Psi_{nl}^{(Ray)}$ and $\Psi_{nl}^{(Aer)}$ comprising up to six independent coefficients (see Appendix A.1 for more details). For Rayleigh scattering, $\Psi_l^{(Ray)}$ coefficients are given in Table 2; for aerosol scattering, they may be generated through electromagnetic scattering calculations.

Linearized properties. We provide two examples of linearized-IOP inputs here for purposes of illustration. For more details on linearized optical properties, the reader is referred to [R1]; see also the treatment in Spurr and Christi (2014). If we require profile Jacobians with respect to the set $\{\chi_n\}$, then it is easy to write down linearized optical properties corresponding to Eq. (3.1) above:

$$\frac{\partial \Delta_n}{\partial \chi_n} = \alpha_n; \quad \frac{\partial \omega_n}{\partial \chi_n} = -\frac{\omega_n \alpha_n}{\Delta_n}; \quad \frac{\partial \beta_{nl}}{\partial \chi_n} = 0. \tag{3.2}$$

Again if we now suppose that layer columns $\{\chi_n(\xi)\}$ are known functions of some bulk property ξ, then we can define linearized inputs required for a calculation of a bulk property Jacobian with respect to ξ:

$$\frac{\partial \Delta_n}{\partial \xi} = \frac{\partial \chi_n}{\partial \xi} \alpha_n; \quad \frac{\partial \omega_n}{\partial \xi} = -\frac{\omega_n}{\Delta_n} \alpha_n; \quad \frac{\partial \beta_{nl}}{\partial \xi} = 0. \tag{3.3}$$

These are pure absorption cases with no derivatives of the expansion coefficients.

We give one more example, an important one. We assume that the analytic dependencies of air column amounts $\{a_n\}$ and gas column amounts $\{\chi_n\}$ on layer temperatures $\{T_n\}$ are known, and that the trace gas cross-sections $\alpha_n(T_n)$ have known temperature dependence. In other words, we know the derivatives $\partial a_n / \partial T_n$, $\partial \chi_n / \partial T_n$ and $\partial \alpha_n / \partial T_n$. We then write down temperature-Jacobian linearized IOPs:

$$\frac{\partial \Delta_n}{\partial T_n} = \frac{\partial G_n}{\partial T_n} + \frac{\partial R_n}{\partial T_n}; \quad \frac{\partial \omega_n}{\partial T_n} = \frac{\frac{\partial R_n}{\partial T_n} - \omega_n \frac{\partial \Delta_n}{\partial T_n}}{\Delta_n}; \quad \frac{\partial \beta_{nl}}{\partial T_n} = \frac{\beta_{nl}^{(Ray)} - \beta_{nl}}{R_n + s_n} \frac{\partial R_n}{\partial T_n}; \tag{3.4a}$$

$$\frac{\partial G_n}{\partial T_n} = \chi_n \frac{\partial \alpha_n}{\partial T_n} + \alpha_n \frac{\partial \chi_n}{\partial T_n}; \quad \frac{\partial R_n}{\partial T_n} = \sigma_{Ray} \frac{\partial a_n}{\partial T_n}. \tag{3.4b}$$

In Spurr and Christi (2014), linearized optical inputs were defined with regard to a set of parameters which characterize certain atmospheric constituent distributions. The canonical example here is a Gaussian-shaped aerosol or trace-gas plume parameterized by the total aerosol or trace-gas optical depth τ_0, the Gaussian peak height z_p and the peak half width h_w. Given a discretized plume profile of aerosol or trace gas optical depth $\{\epsilon_n\}$ parameterized in this manner one can write down optical property derivatives of $\{\epsilon_n\}$ with respect to these three plume characteristics $\{\tau_0, z_p, h_w\}$, and hence obtain linearized IOP inputs for a LIDORT or VLIDORT calculation of Jacobians with respect to these parameters.

3.2 Considerations for Jacobian Outputs

For reasons of numerical stability, it is preferable to give linearized IOPs in normalized form, that is, we specify $\xi \frac{\partial \Delta_n}{\partial \xi}$ instead of $\frac{\partial \Delta_n}{\partial \xi}$, for example. The corresponding weighting function output will be normalized in the same manner. This consideration does not apply to surface property Jacobians; inputs for these are always unnormalized, as are the resulting surface-property Jacobians themselves. [This latter consideration is important, as it is quite possible to have a dark surface (zero albedo) with a non-zero albedo Jacobian].

Weighting function outputs are often related, particularly if linearized IOPs show proportionality. Thus for example, it is possible to derive an additional profile Jacobian for a second trace gas absorber by multiplying the equivalent Jacobian for the first absorber by the ratio of the absorption optical depths for the two gases. This example is discussed in Van Oss and Spurr (2002), and other authors have looked at defining a basic set of optical property weighting functions from which all physical Jacobians can be derived [see for example Ustinov (2005)].

In this regard, we look next at the generation of Jacobian output at Level boundaries.

Layers-to-Levels Jacobian Output.

The (V)LIDORT models ingest *layer* IOPs and their derivatives, and profile Jacobian outputs are for layer-defined quantities. In Spurr and Christi (2014), it was shown that profile Jacobians with respect to *level quantities* (that is, constituent profiles specified at layer boundaries, for example volume mixing ratios for trace species) can be determined from (V)LIDORT Jacobian output specified with respect to the *layer* optical depths. An example with molecular absorbers will suffice to illustrate this conversion.

Suppose we have an atmosphere in which quantities $\{p_n, T_n, h_n, v_n\}$ (pressure, temperature, height, trace-gas volume mixing ratios) are defined at levels $n = 0, 1, \ldots N$. The air density at each level is $\rho_n = Cp_n/T_n$ with the constant of proportionality C expressed in terms of the density at standard temperature and pressure (STP). In a simple application of the hydrostatic equation, we can write the following for the absorber and air column densities G_n and R_n:

$$G_n = \frac{1}{2}Cd_n\left\{\frac{p_{n-1}v_{n-1}\alpha_{n-1}}{T_{n-1}} + \frac{p_nv_n\alpha_n}{T_n}\right\}; \quad R_n = \frac{1}{2}Cd_n\sigma_{Ray}\left\{\frac{p_{n-1}}{T_{n-1}} + \frac{p_n}{T_n}\right\}. \quad (3.5)$$

Here, $d_n = h_{n-1} - h_n$ is the layer height difference, and α_n the level cross sections (which may be functions of pressures and temperatures at these levels). We can write derivatives of these layer optical depths with respect to level VMRs or temperatures for levels $k = n$ or $k = n - 1$:

$$\frac{\partial G_n}{\partial v_k} = \frac{1}{2}Cd_n\frac{p_k\alpha_k}{T_k}; \quad \frac{\partial R_n}{\partial v_k} = 0. \quad (3.6)$$

Given the layer Jacobian output $\frac{\partial I}{\partial G_n}$ and $\frac{\partial I}{\partial R_n}$ obtained in the usual way from LIDORT, we obtain level VMR Jacobians by means of a chain-rule transformation:

$$\frac{\partial I}{\partial v_0} = \frac{\partial I}{\partial G_1}\frac{\partial G_1}{\partial v_0}; \quad (3.7a)$$

$$\frac{\partial I}{\partial v_k} = \frac{\partial I}{\partial G_k}\frac{\partial G_k}{\partial v_k} + \frac{\partial I}{\partial G_{k+1}}\frac{\partial G_{k+1}}{\partial v_k}, \quad k = 1, \ldots N - 1; \quad (3.7b)$$

$$\frac{\partial I}{\partial v_N} = \frac{\partial I}{\partial G_N}\frac{\partial G_N}{\partial v_N}. \quad (3.7c)$$

Level temperature Jacobians $\frac{\partial I}{\partial T_k}$ may be obtained in a similar manner to those in Eqs. (3.7), given starting points obtained by differentiating Eq. (3.5):

$$\frac{\partial G_n}{\partial T_k} = \frac{1}{2}Cd_n\frac{p_kv_k}{T_k}\left\{\frac{\partial \alpha_k}{\partial T_k} - \frac{\alpha_k}{T_k}\right\}; \quad \frac{\partial R_n}{\partial T_k} = -\frac{1}{2}Cd_n\frac{p_k\sigma_{Ray}}{T_k^2}. \quad (3.8)$$

Level Jacobians $\frac{\partial I}{\partial T_k}$ will include both sets of *layer* Jacobians $\frac{\partial I}{\partial G_n}$ and $\frac{\partial I}{\partial R_n}$ from LIDORT or VLIDORT. More details can be found in the above reference.

3.3   Temperature Jacobians in the Thermal Regime

In the thermal regime, we have additional temperature dependence arising through the Planck function inputs $B_n(\nu) = \frac{2h\nu^3}{c}\left[\frac{1}{\exp\frac{h\nu}{ZT_n} - 1}\right]$, $n = 0, 1, \ldots N$ (level index). Here ν is the frequency, c and h the speed of light and Planck's constant respectively, and

Z is Boltzmann's constant. While it is easy to obtain $\frac{\partial B_n(\nu)}{\partial T_n}$ by explicit differentiation of the Planck function, we need additional (V)LIDORT derivatives with respect to $B_n(\nu)$ in order to obtain the complete temperature Jacobian in the thermal regime:

$$\frac{\partial I}{\partial T_n} = \left(\frac{\partial I}{\partial T_n}\right)_{optical} + \frac{\partial I}{\partial B_n}\frac{\partial B_n}{\partial T_n}. \qquad (3.9)$$

The first term in this equation is the level-Jacobian result obtained with linearized IOP inputs, as in Eqs. (3.7) and (3.8). For the second term in Eq. (3.9), the computation of $\frac{\partial I}{\partial B_n}$ is not trivial in a multiple-scattering multi-layer atmosphere, and this feature was only introduced recently into LIDORT (in 2015) and VLIDORT (2016). Here we outline the solution for the LIDORT case.

We start with Eqs. (2.12a) and (2.12b), in which the thermal source particular integral in layer n is expressed as a linear combination of Planck functions B_n and B_{n-1} at the layer boundaries. Differentiation with respect to these quantities is then straightforward.

Given the boundary-value problem $\mathbb{A}\mathbb{C} = \mathbb{B}$ and its solution $\mathbb{C} = \mathbb{A}^{-1}\mathbb{B}$, the linearization of this is $\delta_k\mathbb{C} = \mathbb{A}^{-1}\delta_k\mathbb{B}$, where we have used the shorthand $\delta_k = \frac{\partial}{\partial B_k}$. Matrix $\mathbb{A}$ has no dependence on the Planck functions ($\delta_k\mathbb{A} = \mathbf{0}$), and $\delta_k\mathbb{B}$ may be written down by an explicit examination of the entries in Eq. (2.12a) and their linearizations. Indeed, for the first two levels:

$$\delta_0\mathbb{B} = \begin{pmatrix} -\mathbf{P}_{1\alpha}^+ \\ -\mathbf{R}_{1\alpha}^+ \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}; \quad \delta_1\mathbb{B} = \begin{pmatrix} -\mathbf{Q}_{1\alpha}^+ \\ \mathbf{P}_{2\alpha}^\pm - \mathbf{S}_{1\alpha}^\pm \\ -\mathbf{R}_{2\alpha}^+ \\ 0 \\ \cdots \\ 0 \end{pmatrix} ., \; \text{etc} \ldots \qquad (3.10)$$

The trick here is to make sure that the vector entries $\delta_k\mathbb{B}$ are filled up correctly—a variable defined at level boundary k will have an effect on adjacent layers above and below this boundary. Since we have the inverse matrix $\mathbb{A}^{-1}$ from the original solution of the radiance field, the linearized vector $\delta_k\mathbb{C}$ is readily computed.

The linearized discrete ordinate solution in layer n with respect to the Planck function B_k at level k can now be written

$$\frac{\partial \mathbf{I}_n^\pm(x)}{\partial B_k} = \sum_{\alpha=1}^N \left[\frac{\partial L_{n\alpha}}{\partial B_k}\mathbf{X}_{n\alpha}^\pm e^{-xk_{n\alpha}} + \frac{\partial M_{n\alpha}}{\partial B_k}e^{-(\Delta_n-x)k_{n\alpha}}\mathbf{X}_{n\alpha}^\mp\right] + \frac{\partial \mathbf{G}_n^\pm(x)}{\partial B_k}. \qquad (3.11)$$

Here the derivatives $\left\{\frac{\partial L_{n\alpha}}{\partial B_k}, \frac{\partial M_{n\alpha}}{\partial B_k}\right\}$ emerge from the $\delta_k\mathbb{B}$ vector, while $\frac{\partial \mathbf{G}_n^\pm(x)}{\partial B_k}$ follows from explicit differentiation of Eqs. (2.12a) and (2.12b). Derivatives $\frac{\partial \mathbf{G}_n^\pm(x)}{\partial B_k}$ exist only for $n = k$ or $n = k + 1$, but $\left\{\frac{\partial L_{n\alpha}}{\partial B_k}, \frac{\partial M_{n\alpha}}{\partial B_k}\right\}$ are defined for all values of n.

Surface emission Jacobian. We make one remark here about Jacobians with respect to the surface Planck function B_{surf}. Clearly none of the Green's function particular integrals in Eqs. (2.12) depend on this quantity; we are left with a linearization of the boundary value problem. If we write $\delta_{surf} = \frac{\partial}{\partial B_{surf}}$, the linearized problem is $\delta_{surf}\mathbb{C} = \mathbb{A}^{-1}\delta_{surf}\mathbb{B}$, where from Eq. (2.13), we find that the only surviving term in the differentiation is:

$$\delta_{surf}\mathbb{B} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \mathbf{E} \end{pmatrix}. \tag{3.12}$$

Linearization of the post-processed solutions with respect to B_{surf} may be determined readily enough through differentiation of the components of the post-processed radiance solution in Eqs. (2.16).

3.4 Pseudo-spherical Capability, Delta-M Scaling, N-T Ansatz

Pseudo-spherical capability

LIDORT and VLIDORT can deal with solar-beam attenuation in a curved spherical-shell atmosphere. This is the pseudo-spherical (P-S) approximation. We use the "average-secant" parameterization (Spurr 2002), in which the solar beam transmittance in layer n is expressed as an exponential function of vertical optical thickness x in that layer:

$$\Theta_n(x) = T_{n-1}e^{-\lambda_n x}; \quad \lambda_n = \frac{1}{\Delta_n}\ln\frac{T_{n-1}}{T_n}; \quad T_n = \exp\left[-\sum_{k=1}^{n}f_{nk}\Delta_k\right]. \tag{3.13}$$

Here, T_{n-1} and T_n are the ray-traced transmittances to the layer upper and lower boundaries respectively, and Δ_n is the vertical optical thickness of the layer. In a plane parallel atmosphere, $\lambda_n \equiv 1/\mu_0$, where $-\mu_0$ is the SZA cosine. Quantities $\{T_n\}$ are expressed in terms of the ray-traced Chapman factors f_{nk}, purely geometrical quantities defined as the ratios of layer slant-path distances to vertical height differences. For a geometrical (non-refracting) atmosphere, one need only specify the height grid $\{h_n\}, n = 0, 1 \dots N$, and the earth radius R_{Earth} as inputs to the model.

Linearization of the pseudo-spherical scheme is straightforward. Without loss of generality, we can define derivatives with respect to the set $\{\Delta_n\}$—other linearizations can be obtained through chain-rule differentiation. From Eq. (3.13), it follows that:

$$\frac{\partial \mathrm{T}_n}{\partial \Delta_k} = -f_{nk} \mathrm{T}_n; \quad \frac{\partial \lambda_n}{\partial \Delta_k} = \frac{f_{nk} - f_{n-1,k} - \delta_{nk} \lambda_n}{\Delta_n}. \tag{3.14}$$

Here δ_{nk} is the Kronecker-delta; note that both these quantities have cross-layer derivatives.

Delta-M Scaling and the N-T Ansatz

The delta-M scaling procedure (Wiscombe 1977) is a useful approximation for dealing with sharply-peaked forward scattering. The delta-M scaling is standard to LIDORT and VLIDORT.

In scalar RT with LIDORT, the delta-M *ansatz* is to approximate the true phase function in a given layer n as $\Phi_n(\Theta) \approx 2\pi f_n \delta(1 - \cos \Theta) + (1 - f_n)\tilde{\Phi}_n(\Theta)$ in terms of delta-function extinction in the forward scattering direction, and a smoother "residual" phase function $\tilde{\Phi}_n(\Theta)$. Introducing this approximation into the RTE allows use of a set of "scaled" IOPs determined by selection of the truncation factor f_n. This procedure is performed for each layer, and for any discrete-ordinate RT calculation with N_d streams in the half-space, the relevant formulae for scaled IOPs are:

$$f_n = \frac{\beta_{nM}}{(2M+1)}; \tilde{\Delta}_n = \Delta_n(1 - f_n \omega_n); \tilde{\omega}_n = \frac{\omega_n(1 - f_n)}{(1 - f_n \omega_n)}; \tilde{\beta}_{nl} = \frac{\beta_{nl} - f_n(2l+1)}{(1 - f_n)}. \tag{3.15}$$

Here $M = 2N_d$, and $l = 0, 1, \ldots 2N_d - 1$ in the definition of $\tilde{\beta}_{nl}$.

In VLIDORT, we take the truncation factor f_n as defined Eq. (3.15), and adopt the following scaling for the set of Greek coefficients: 4 coefficients $\{\alpha_{nl}, \beta_{nl}, \delta_{nl}, \zeta_{nl}\}$ will scale as β_{nl} in Eq. (3.15), while γ_{nl} and ε_{nl} will scale as $\tilde{\gamma}_{nl} = \gamma_{nl}/(1 - f_n)$. This is the system adopted in Chami et al. (2001), where more details may be found.

Linearized "delta-M scaled" IOPs for LIDORT were first discussed in Van Oss and Spurr (2002). If we suppose that unscaled properties $\{\Delta_n, \omega_n, \beta_{nl}\}$ in layer n have known derivatives $\{\dot{\Delta}_n, \dot{\omega}_n, \dot{\beta}_{nl}\}$ with respect to some quantity ξ varying in that layer, then the linearized scaled IOPs are (we also include the VLIDORT γ_{nl} scaling for completenes):

$$\dot{f}_n = \frac{\dot{\beta}_{nM}}{(2M+1)}; \dot{\tilde{\Delta}}_n = \dot{\Delta}_n(1 - f_n \omega_n) - \Delta_n(\dot{f}_n \omega_n + f_n \dot{\omega}_n); \tag{3.16a}$$

$$\dot{\tilde{\omega}}_n = \frac{\tilde{\omega}_n(\dot{f}_n \omega_n + f_n \dot{\omega}_n) + \dot{\omega}_n(1 - f_n) - \omega_n \dot{f}_n}{(1 - f_n \omega_n)}; \tag{3.16b}$$

$$\dot{\tilde{\beta}}_{nl} = \frac{\dot{\beta}_{nl} - \dot{f}_n(2l+1) + \dot{f}_n \tilde{\beta}_{nl}}{(1 - f_n)}; \quad \dot{\tilde{\gamma}}_{nl} = \frac{\dot{\gamma}_{nl} + \dot{f}_n \tilde{\gamma}_{nl}}{(1 - f_n)}. \tag{3.16c}$$

The *Nakajima and Tanaka* [N-T] ansatz (Nakajima and Tanaka 1988) is a well-known procedure for obtaining more accurate RT solutions when the delta-M scaling is in force. The N-T process works by replacing the internal discrete-ordinate single-scatter (SS) calculation (which is based on the use of a set of $2N_d$ phase

function/matrix expansion coefficients) with an "exact" single scatter calculation using the phase function itself (LIDORT) or the phase matrix (VLIDORT). When the delta-M scaling applies, the exact single scatter computation is performed using the *unscaled* phase function/matrix and single scatter albedo, but using the *scaled* optical thickness variable.

The complete NT single-scatter calculation was first described for LIDORT and VLIDORT in [R1], including linearization aspects. In earlier versions of the SS calculations, phase functions or F-matrices were always computed internally using input sets of spherical-function expansion coefficients and internal calculations of Legendre or spherical-function polynomials.

There have been two changes in the implementation of SS computations in recent versions of LIDORT and VLIDORT. Now, it is possible to ingest phase function (LIDORT) or F-matrix (VLIDORT) inputs *directly* without the need for reconstructions using expansion coefficients. In addition, the entire SS calculation has now been modularized, so that it can be used either as a single call from LIDORT or VLIDORT, or as a stand-alone ***First-Order*** (FO) model for use in other applications. This FO model is described in Sect. 4.4.

4 (V)LIDORT Supplements and First-Order Codes

4.1 Overview

From an early stage, the bidirectional reflectance distribution function (BRDF) surface capability has always been present in both the scalar and vector RT codes (Spurr 2004). In earlier versions of the codes, configuration control for BRDF inputs was an integral part of the overall input control scheme for the models. However, following user feedback after the first LSR review [R1], it was decided to move the computation of overall BRDF properties into a separate supplementary module, so that the main RT code is not encumbered by choices of BRDF kernels and their specifications. This not only simplifies the (V)LIDORT input configurations, but also allows the user much more flexibility with setting up BRDF conditions.

The motivation here follows that for the atmospheric optical property inputs, namely, that (V)LIDORT only ingests the *total* optical layer quantities (optical thicknesses, single scattering albedos, scattering matrices/phase functions); the RT code does not distinguish between individual trace gas absorbers, or between Rayleigh and particulate scattering. Similarly with the BRDF surface: (V)LIDORT requires only the *total* BRDF inputs for direct-bounce and diffuse-field reflectance—the code no longer distinguishes individual BRDF kernels that make up the total reflectance.

The BRDF supplement computation is the subject of Sect. 4.2 below. This follows on from the treatment in [R1], though in recent years a number of additional BRDF kernels have been introduced. We also discuss the BRDF supplement input/output,

and give an example of the supplement usage. BRDF kernels are detailed in Appendix B.

The surface-leaving (SL) supplement was introduced in 2011 to the LIDORT and VLIDORT codes, and postdates the [R1] review. For land surfaces, there is now a facility for SIF (Solar-induced Fluorescence) radiances in the NIR—this is based on a recent climatology which generates a double-Gaussian isotropic fluorescence field depending only on the SZA, and time geolocation (Frankenberg et al. 2012). For water-leaving radiances, there is now a more sophisticated ocean optics model, and an extension to non-isotropy through rough-surface glitter. Surface-leaving options are discussed in Sect. 4.3.

Although it is possible for both RT codes to ingest pre-calculated "first-order" (FO) radiation (single scattering and direct-bounce intensity/Stokes-vector fields, direct thermal and surface emission), there is no specific supplement that will compute these fields. However, the FO code Version 1.5, which is now part of the (V)LIDORT models, can be thought of as a supplement in its own right—all FO subroutines are stand-alone, and are normally called inside the (V)LIDORT master modules when a single Boolean flag controlling their execution is set. The FO codes are described in Sect. 4.4 below.

In regard to the FO calculations, a third supplement has been newly written for Versions 2.8 (VLIDORT) and also for 3.8 (LIDORT), following feedback from a number of users. These new adjuncts are called the F-Matrix supplement "FMATRIX" (for VLIDORT) and the phase function supplement "PHASFUNC" (for LIDORT); they are designed to ingest *external* data sets of scattering matrices (vector code) or phase functions (scalar code) and from these, generate optical property input for VLIDORT and LIDORT in two forms—first, for accurate single-scattering computations, a set of scattering matrices or phase functions at user-defined scattering angles, and second, all necessary spherical function expansion coefficients for use in the LIDORT and VLIDORT multiple scatter computations. These supplements are discussed in Sect. 4.5.

4.2 Bidirectional Reflectance Supplement

4.2.1 BRDF Requirements for the RT Models

We confine attention here to the scalar LIDORT surface boundary condition linking reflected upwelling with incident downwelling radiances, which may be written:

$$I^{\uparrow}(\mu, \varphi) = \int_0^{2\pi} \int_0^1 \rho(\mu, \mu', \varphi - \varphi') I^{\downarrow}(\mu, \varphi) \mu' d\mu' d\varphi'. \tag{4.1}$$

We have assumed that the BRDF $\rho(\mu, \mu', \varphi - \varphi')$ depends only on the relative azimuth direction. For the direct-bounce condition (single reflection of atten-

uated direct sunlight in a spherically curved atmosphere), we require BRDFs $\rho(\mu_1, \mu_0, \varphi_1 - \varphi_0)$ for incoming SZA cosines μ_0, outgoing line-of-sight cosines μ_1 and relative azimuth angles $\varphi_1 - \varphi_0$.

We express the BRDF as a Fourier cosine series in relative azimuth:

$$\rho(\mu, \mu', \varphi - \varphi') = \sum_{m=0}^{2N_d-1} (2 - \delta_{m0}) \rho_m(\mu, \mu') \cos m(\varphi - \varphi'). \qquad (4.2)$$

For surface reflectance of the diffuse field, we then require the four Fourier-series components $\rho_m(\mu_i, \mu_0)$, $\rho_m(\mu_i, \mu_j)$, $\rho_m(\mu_1, \mu_j)$ and $\rho_m(\mu_1, \mu_0)$. Here, $\{\mu_j\}$, $j = 1 \cdots N_d$ are the N_d discrete ordinate polar streams. Note that the cosine series for $\rho_m(\mu_1, \mu_0)$ components represents a truncated form of the direct-bounce intensity, and these components are not normally required when the direct-bounce BRDF is computed.

LIDORT requires these five _total_ BRDF quantities (direct-bounce BRDF, four Fourier-component forms) to specify completely the surface boundary condition. The task for the BRDF supplement is to provide these total inputs.

4.2.2   BRDF Supplement: Kernel Functions

Some of the material in this section has appeared before in [R1], but we include it for completeness. Details on the kernels may also be found in Appendix B.

The BRDF supplement relies on the combination of one or more analytic semi-empirical BRDF kernel functions. Thus, the _scalar_ total BRDF $\rho_{total}(\mu, \mu', \varphi - \varphi')$ is specified as a linear weighting of (up to) four semi-empirical kernel functions:

$$\rho_{total}(\mu, \mu', \varphi - \varphi') = \sum_{k=1}^{4} R_k \rho_k(\mu, \mu', \varphi - \varphi', \mathbf{b}_k). \qquad (4.3)$$

Here, R_k are linear combination coefficients or "kernel amplitudes". For each kernel, the geometrical dependence is known, but the kernel function will (in some but not all cases) depend on the values taken by a vector $\mathbf{b}_k$ of pre-specified parameters. Well-known examples include the Cox-Munk BRDF for glitter reflectance from the ocean (Cox and Munk 1954a, b); and the Ross-Li portfolio of land-surface kernels used in the MODIS BRDF retrieval system (Lucht and Roujean 2000; Lucht et al. 2000). The Cox-Munk kernel comprises a wave-facet probability distribution function (depending on wind-speed W), and a Fresnel reflection function (depending on the air-water relative refractive index n_{rel}). In this case, vector $\mathbf{b}_k = \{W, n_{rel}\}$ has two elements. For a Lambertian surface, there is only one kernel: $\rho_{Lambertian}(\mu, \mu', \varphi - \varphi') = 1$ for all incident and reflected angles, and coefficient R_1 is just the Lambertian albedo.

Altogether, the LIDORT BRDF supplement has 13 kernel choices, and these are summarized in Table 3, along with associated non-linear parameters. A full

Table 3 The BRDF Kernel functions in LIDORT

Index	Name	Size $\mathbf{b}_k$	Reference
1	Lambertian	0	–
2	Ross thick	0	Wanner et al. (1995)
3	Ross thin	0	Wanner et al. (1995)
4	Li sparse	2	Wanner et al. (1995)
5	Li dense	2	Wanner et al. (1995)
6	Roujean	0	Wanner et al. (1995)
7	Hapke	3	Hapke (1993)
8	Rahman	3	Rahman et al. (1993)
9	Cox-Munk	2	Cox and Munk (1954a)
10	BPDF soil	1	Maignan et al. (2009)
11	BPDF vegetation	1	Maignan et al. (2009)
12	BPDF NDVI	3	Maignan et al. (2009)
13	New Cox-Munk	3	Sayer et al. (2010)

discussion of the first 9 kernels is in Spurr (2004). Kernels 10–12 are scalar versions of polarized land-surface reflectances, based on code provided by François-Marie Bréon (Maignan et al. 2009). Kernel 13 is based on new water-leaving and ocean reflectance parameterizations provided in the context of the 6S model (Kotchenova et al. 2006; Andrew Sayer, private communication).

For land surfaces, analytic kernel models have trouble dealing with distinctive vegetation signatures such as "hot-spots", and in an interesting new departure, Huang et al. (2017) have used a detailed 3-D canopy-vegetation model to generate look-up tables of VLIDORT BRDF inputs in place of the usual kernel-driven supplement calculation Also in this regard, a snow surface data-base (Lyapustin et al. 2010) has been adapted for use in VLIDORT.

Fourier components of the kernel BRDFs are calculated through:

$$\rho_{km}\big(\mu, \mu', \mathbf{b}_k\big) = \frac{1}{2\pi} \int\limits_0^{2\pi} \rho_k\big(\mu, \mu', \varphi, \mathbf{b}_k\big) \cos m\varphi \, d\varphi. \tag{4.4}$$

The azimuth integration done by double numerical quadrature over the ranges $[0, \pi]$ and $[-\pi, 0]$; the number of BRDF azimuth quadrature abscissa is set to 100 to obtain a numerical accuracy of 10^{-4} for all kernels considered (Spurr 2004).

Linearization of this BRDF scheme was reported in Spurr (2004), and a mechanism developed for the generation of surface property weighting functions with respect to the kernel amplitudes R_k and to elements b_{pk} of the kernel parameters $\mathbf{b}_k$. The complete discrete ordinate RT model is differentiable with respect to these surface properties, given the following kernel derivatives:

$$\frac{\partial \rho_{total}(\mu, \mu', \varphi)}{\partial R_k} = \rho_k(\mu, \mu', \varphi, \mathbf{b}_k); \quad \frac{\partial \rho_{total}(\mu, \mu', \varphi)}{\partial b_{pk}} = \frac{\partial \rho_k(\mu, \mu', \varphi, \mathbf{b}_k)}{\partial b_{pk}}.$$

(4.5)

The amplitude derivative is straightforward. The (non-linear) parameter derivatives depend on the empirical formulation of the kernel in question; all kernels in the BRDF supplement are analytically differentiable with respect to any such parameter dependencies.

Remark In the vector code VLIDORT, the BRDF is actually a 4×4 matrix, linking incident and reflected Stokes 4-vectors. The scalar BRDF scheme outlined above has been fully implemented in VLIDORT by setting the $\{1,1\}$ element of a 4×4 vector kernel P_k equal to the corresponding scalar kernel function ρ_k; all other BRDF matrix elements are then zero.

4.2.3   White-Sky and Black-Sky Albedos

It is sometimes useful to normalize the total BRDF inputs according to a choice of spherical albedo—either the total spherical or *white-sky albedo* (WSA), or the directional *black-sky albedo* (BSA), the latter dependent on the solar zenith angle (θ_0). This is particularly appropriate when using the MODIS-based system of semi-empirical BRDF kernels, where it is sometimes necessary to normalize according to an external value of the albedo.

Assuming the kernel BRDFs to be normalized to $1/\pi$, the two albedos are defined by:

$$A_{WSA} = 4 \int_0^1 \int_0^1 \mu \mu' \rho_0(\mu, \mu') d\mu d\mu'; \quad A_{BSA}(\mu_0) = 2 \int_0^1 \mu' \rho_0(\mu', \mu_0) d\mu'.$$

(4.6)

Here $\mu_0 = \cos \theta_0$, and $\rho_0(\mu, \mu')$ is the Fourier-zero component of the total BRDF. Assuming a 4-kernel total BRDF, we can write the albedo-scaled (normalized) BRDF as

$$\tilde{\rho}(\mu, \mu', \varphi - \varphi') = \frac{\tilde{A}}{A_{SA}} \sum_{k=1}^4 R^{(k)} \rho^{(k)}(\mu, \mu', \varphi - \varphi').$$

(4.7)

Here, $\tilde{A}$ is a given _external_ spherical albedo (either WSA or BSA), and the _internal_ spherical albedo is $A_{SA} = \sum_{k=1}^4 R^{(k)} A_{SA}^{(k)}$, where $A_{SA}^{(k)} = A_{WSA}^{(k)}$ for the white-sky case, or $A_{SA}^{(k)} = A_{BSA}^{(k)}(\mu_0)$ for the black-sky case. In order to obtain values of $A_{SA}^{(k)}$, the half-space polar integrals in Eq. (4.6) are done by Gaussian quadrature using abscissa and weights $\{\mu_p, w_p\}$, $p = 1 \ldots N_p$. The default value of N_p is currently 24. We

thus require computation of the Fourier components $\rho_0^{(k)}(\mu_0, \mu_p)$ (BSA case), or $\rho_0^{(k)}(\mu_q, \mu_p)$ (WSA case). The BRDF supplement has a consistency check on the magnitude of the total internally-calculated BRDF spherical albedo A_{SA}; this should be non-negative and take a value in the interval [0, 1].

Remark Albedo-scaling (if selected) applies to all BRDF Fourier components as well as the direct bounce term. In the BSA case, albedo-scaling is dependent on the SZA through μ_0, and it then follows that the <u>scaled</u> components $\tilde{\rho}_m(\mu_j, \mu_i)$ and $\tilde{\rho}_m(\mu_j, \mu_1)$ will pick up a dependence on θ_0 which the original unscaled components did not possess. Because of this additional dependence, BSA scaling can only be applied to all BRDF outputs for a single value of θ_0.

WSA/BSA Linearizations

Starting with analytic partial derivatives in Eq. (4.5), application of albedo scaling requires additional differentiation. Indeed, from Eq. (4.7) we find (dropping the geometrical variables for convenience):

$$\frac{\partial \tilde{\rho}}{\partial R^{(k)}} = \frac{\tilde{A}\rho^{(k)} - \tilde{\rho}A_{SA}^{(k)}}{A_{SA}}; \quad \frac{\partial \tilde{\rho}}{\partial b^{(k)}} = \frac{\tilde{A}R^{(k)}\frac{\partial \rho^{(k)}}{\partial b^{(k)}} - \tilde{\rho}\frac{\partial A_{SA}^{(k)}}{\partial b^{(k)}}}{A_{SA}}. \tag{4.8}$$

This last result requires the following computations to be added to the BRDF supplement:

$$\frac{\partial A_{WSA}^{(k)}}{\partial b^{(k)}} = 4 \int_0^0 \int_0^1 \mu\mu' \frac{\partial \rho_0^{(k)}}{\partial b^{(k)}}(\mu, \mu')d\mu d\mu'; \quad \frac{\partial A_{BSA}^{(k)}}{\partial b^{(k)}}(\mu_0) = 2 \int_0^1 \mu' \frac{\partial \rho_0^{(k)}(\mu', \mu_0)}{\partial b^{(k)}}d\mu'. \tag{4.9}$$

These derivatives are also determined using quadrature $\{\mu_p, w_p\}$. It is possible to generate derivatives of the scaled BRDFs with respect to the user-supplied external spherical albedo $\tilde{A}$. The WSA/BSA-derivative is easy to write down; indeed, from Eq. (4.6) we find that $\frac{\partial \tilde{\rho}}{\partial \tilde{A}} = \frac{\tilde{\rho}}{\tilde{A}}$.

For the vector BRDF supplement, spherical albedos are derived from the (1,1) element of the (Fourier-zero) component of the BRDF matrix. All other elements of this matrix are scaled by the same spherical albedo.

4.3 Surface Leaving Supplement

The SL supplement is a separate system of (V)LIDORT-based software that generates a source of upwelling radiance at the lower boundary, for instance, "water-leaving" radiance from the ocean, or near-infrared solar-induced fluorescence (SIF) from vegetation. The SL contributions to (V)LIDORT from this supplement will depend

on the downwelling solar illumination angle θ_0, By analogy with the BRDF case, we need to specify a "direct" source of upwelling radiance, depending on incoming and outgoing directions, and for the diffuse field, Fourier components of the surface leaving radiance. The three SL terms are (restricting attention to the scalar case):

$$S_m(\mu_i, \mu_0), \ i = 1, \ldots, N_d; \quad S_m(\mu, \mu_0); \quad S_{direct}(\mu, \mu_0, \varphi - \varphi_0). \quad (4.10)$$

Here, $m = 0, 1, \ldots, 2N_d - 1$ is the Fourier component index as before. The first term is the (Fourier component of) upwelling radiance into the polar discrete ordinate directions $\{\mu_i\}$, and this is required for the diffuse-field boundary condition at the lower surface. The second term is the (Fourier component of) upwelling radiance into user-defined stream direction μ, and this is required for post-processing of the discrete ordinate solution (source function integration). In the post-processing, it is more accurate to use the complete term $S_{Direct}(\mu, \mu_0, \varphi - \varphi_0)$ in place of the Fourier-series truncation. This "correction" is analogous to the precise calculation of the direct-beam surface reflection based on the values of $S_m(\mu_0, \mu)$, and in this case, the Fourier terms $S_m(\mu_0, \mu)$ are not needed in the (V)LIDORT calculations.

The simplest SL specification comprises an isotropic term $S^*(\mu_0)$ which depends only on the SZA. In this case, $S_m(\mu_i, \mu_0) = 0, m \geq 1$ and $S_0(\mu_i, \mu_0) = S^*(\mu_0)$ for all outgoing polar directions μ_i, and also $S_{Direct}(\mu, \mu_0, \varphi - \varphi_0) = S^*(\mu_0)$.

Linearization. We assume that surface-leaving terms depend only on quantities intrinsic to the surface (no atmospheric dependence). We require the SL supplement to supply partial derivatives of terms in Eq. (4.10) with respect to some surface property ξ (which might be the wind speed or the fluorescence at 755 nm; see below); the main RT codes are then able to ingest these linearized inputs, thereby making it possible for (V)LIDORT to produce Jacobian output with respect to surface property ξ. We now discuss the current SL implementations.

Water-leaving. In the first SL supplements to the RT models, water-leaving contributions were regarded as isotropic with no SZA dependence. All water leaving radiances were equal to the flux-normalized (underwater) upwelling radiance S^* in the ocean at the surface (Fresnel transmittance through the surface was taken to be unity). The value of S^* was obtained through an empirical formulation based on marine scattering and absorption (pure water, pigment (chlorophyll) and CDOM). This ocean-optics model was based originally on the specification in the 6S model, but we have updated the model based in part on the more recent work of Morel and Gentili (2009). The ocean-optics model is dependent only on wavelength (in Microns), the salinity and pigment concentration in [mg/M].

The 6S formulation also included a treatment of the ocean-air transmittance factor for a rough surface; this is based on the azimuthal integration (for each solar angle and line of sight angle in the atmosphere) of the reverse-medium glint calculation for the rough-surface interface, taking into account Snell's law of refraction and the conservation of light intensity divided by the square of the reflective index. Glint calculations here are done using the same Cox-Munk calculation as that for the

BRDF (with wind-directionality and refractive index determined through salinity. The whitecap correction to the SL is also required if we are computing water-leaving radiance in conjunction with atmospheric glint with foam.

Although the SLEAVE and BRDF supplements are using the same software for glint reflectance, Fresnel coefficient calculation, plus refractive index and whitecap determination, the supplements have been kept completely separate in the interests of modularity. When using the two supplements together, it is essential that they both operate with a common set of inputs. With this in mind, there is a subroutine which checks the compatibility of SLEAVE and BRDF inputs. Two other "input-check" routines are also present to ensure that geometrical variables are consistent between (V)LIDORT and its BRDF and SL supplements.

Fluorescence. This is based on the double-Gaussian model (Frankenberg et al. 2012) which has now been used in a number of studies on the SIF signature. We would like to thank Dr. Chris O'Dell for allowing us to use this model. The calculation is simple:

$$
S^*(\theta_0) \equiv F(\lambda, \theta_0) = F_{755}(\theta_0) \left\{ A_1 \exp\left[-\frac{(\lambda - \lambda_1)^2}{\sigma_1^2} \right] + A_2 \exp\left[-\frac{(\lambda - \lambda_2)^2}{\sigma_2^2} \right] \right\}.
$$
(4.11)

The wavelengths λ_1 and λ_2 correspond to peaks at 683 nm and 730 nm respectively, and all the Gaussian constants are tabulated in the aforementioned reference. The SIF is isotropic. The fluorescence $F_{755}(\theta_0)$ at 755 nm is based on a large multi-year data set derived from satellite observations, depending on solar angle θ_0, 'epoch' (year, month, day, hour, etc.) and latitude and longitude coordinates. Equation (4.10) is easy to differentiate with respect to the defining parameters. The main interest here is with the amplitude parameter $\xi \equiv F_{755}(\theta_0)$ for which the partial derivative $\partial S^*(\theta_0)/\partial \xi$ is trivial. It is possible to define also Jacobians with respect to the Gaussian parameters in Eq. (4.11), and there is a facility for this option.

4.4 First-Order Codes (Single Scattering, Direct Thermal Calculations)

For the new versions (VLIDORT 2.8, LIDORT 3.8) we have revised and simplified the treatment of single scattering, direct-bounce reflection and direct thermal emission in a curved spherical-shell atmosphere. This section replaces the exposition in [R1].

Dealing first with single scattering in the upwelling direction $\Omega = (\mu, \varphi)$, for a given layer n in an optically-stratified atmospheric model, the vector RTE is (assuming non-refractive geometry):

$$
\frac{d\mathbf{I}(x, \Omega)}{dx} = -\mathbf{I}(x, \Omega) + \frac{\omega_n}{4\pi} A_n(x) \mathbf{\Pi}_n^\uparrow(\Omega, \Omega_0) \mathbf{F}_\odot.
$$
(4.12)

Here, x is the optical thickness coordinate along the line of sight, $\Omega_0 = (-\mu_0, \varphi_0)$ is the solar beam direction, $\mathbf{F}_\odot = (F_\odot, 0, 0, 0)^T$ the solar flux vector, $\mathbf{\Pi}_n^\uparrow(\Omega, \Omega_0)$ the scattering matrix, ω_n the single scattering albedo, and $A_n(x)$ the solar beam attenuation to the point of scatter. Note that, although the directions $\{\Omega, \Omega_0\}$ will change along the line of sight, the angle of scatter remains constant in straight-line geometry.

In the (V)LIDORT models, layers are optically uniform, so ω_n and $\mathbf{\Pi}_n^\uparrow(\Omega, \Omega_0)$ are constant for a given layer n, and $x = \rho_n s$ in terms of the (constant) layer extinction coefficient ρ_n, and path distance s. In terms of the layer vertical height difference h_n and layer vertical optical thickness Δ_n (one of the basic IOP inputs), $\rho_n = \Delta_n / h_n$. Using path distance as the variable (more convenient, as this is independent of wavelength), we integrate along the line-of-sight from $s = 0$ to $s = D_n$ (at the top boundary of the layer):

$$\mathbf{I}(D_n, \Omega) = \mathbf{I}(0, \Omega) e^{-\rho_n D_n} + \mathbf{S}_n^\uparrow; \mathbf{S}_n^\uparrow = \frac{\sigma_n}{4\pi} \mathbf{\Pi}_n^\uparrow(\Omega, \Omega_0) \mathbf{F}_\odot \int_0^{D_n} A_n(s) e^{-\rho_n(D_n - s)} ds.$$

(4.13)

This defines the upwelling layer source term $\mathbf{S}_n^\uparrow$. In addition, we have $\sigma_n = \omega_n \rho_n$ (the scattering coefficient), and the attenuation $A_n(s)$ is found by ray-tracing:

$$A_n(s) = \exp\left[-\sum_{k=1}^n d_{nk}(s) \rho_k \right],$$

(4.14)

where $d_{nk}(s)$ are solar path distances through whole layers $k = 1, \ldots n - 1$ and the partial distance in layer $k = n$ to the point of scatter.

To obtain the upwelling single-scatter field at all levels in the atmosphere, Eq. (4.13) is applied recursively, starting at the surface; if we write $\mathbf{I}_n^\uparrow$ to indicate the upwelling Stokes vector at level boundary n (lower boundary of the layer), then for instance at the top of the atmosphere:

$$\mathbf{I}_0^\uparrow(\Omega) = C_N \mathbf{I}_{Surf}^\uparrow(\Omega) + \sum_{k=1}^N C_{k-1} \mathbf{S}_k^\uparrow(\Omega); \quad C_n = \prod_{k=1}^n e^{-\rho_k D_k}; \quad C_0 = 1. \quad (4.15)$$

Similar expressions pertain for output at other levels. The FO model for the upwelling field does include the direct-bounce surface reflectance of the solar beam:

$$\mathbf{I}_{Surf}^\uparrow(\Omega) = \frac{1}{\pi} A_{Surf} \mu_0 \mathbf{R}(\Omega, \Omega_0) \mathbf{F}_\odot.$$

(4.16)

with A_{Surf} the solar beam attenuation to the surface, and $\mathbf{R}(\Omega, \Omega_0)$ the BRDF reflection matrix at the surface. In the new FO supplement, there is also the capability to generate fields at intermediate points between level boundaries; the main difference

here is that we must now consider partial-layer transmittances and attenuation integrals:

$$\mathbf{I}(s, \Omega) = \mathbf{I}(0, \Omega)e^{-\rho_n s} + \frac{\sigma_n}{4\pi}\mathbf{\Pi}_n^{\uparrow}(\Omega, \Omega_0)\mathbf{F}_\odot \int\limits_0^s A_n(t)e^{-\rho_n(s-t)}dt. \qquad (4.17)$$

The treatment for downwelling radiation is very similar.

In the FO model, there are three options for computing path attenuations. The first option assumes the atmosphere is plane-parallel. In this case we write $s = (h_n - z)/\mu$ in terms of the vertical height difference z from layer-top and the line-of-sight cosine μ, and we have $A_n(s) = T_n \exp[-\rho_n z/\mu_0]$, where T_n is the solar beam transmittance to layer-top. Then we obtain the familiar result:

$$\int\limits_0^{D_n} A_n(s)e^{-\rho_n(D_n-s)}ds = \frac{T_n}{\mu\rho_n}\left[\frac{1 - \exp(-\Delta_n\kappa_n)}{\kappa_n}\right]; \quad \kappa_n = \frac{1}{\mu_0} + \frac{1}{\mu}. \qquad (4.18)$$

The second option allows for curved ray-tracing of the solar beam, but not for the outgoing line-of-sight (the pseudo-spherical approximation). Here we still have $s = (h_n - z)/\mu$, so we could work with vertical coordinates. In the average-secant parameterization used for solar beam attenuation in the multiple-scatter LIDORT and VLIDORT computations, $A_n(s) = T_n \exp[-\lambda_n\rho_n z]$, where the average secant is given by $\lambda_n = \ln[T_n/T_{n+1}]/\rho_n h_n$. In this case, the result in Eq. (4.18) applies, only this time with $\kappa_n = \lambda_n + \frac{1}{\mu}$.

The third choice for the FO model is the fully-spherical situation in which both incoming and outgoing paths are treated for spherical geometry. In this case, the FO model will evaluate attenuation integrals numerically, using a quadrature scheme:

$$\int\limits_0^{D_n} A_n(s)e^{-\rho_n(D_n-s)}ds \cong \sum_{j=1}^{N_q} c_j A_n(s_j)e^{-\rho_n(D_n-s_j)}. \qquad (4.19)$$

Here, the quadrature is $\{s_j, c_j\}$, $j = 1\ldots N_q$ taken over the interval $[0, D_n]$. We have found that Gauss-Legendre integration yields accuracy at the 10^{-6} level in all situations, with N_q varying from 5 to 12. However, there are occasions where care must be taken with this procedure. For an optically thick layer below clear skies, solar beam attenuation through such a layer may become vanishingly small. There is then no light reaching the lowest part of the line-of-sight path, and quadrature over interval $[0, D_n]$ will not work. The remedy is to introduce a cutoff distance D_{nc} such that quadrature in Eq. (4.19) is over the interval $[D_{nc}, D_n]$. The cutoff is determined as follows: assuming that the layer attenuation $\exp[-d_{nn}(s)\rho_n]$ never falls below a certain threshold value A_{crit}, we use ray-tracing and numerical solution methods to solve for D_{nc} through the equation $A_{crit} = \exp[-d_{nn}(D_{nc})\rho_n]$. We have found $A_{crit} = 10^{-8}$ is a suitable threshold.

The situation with direct thermal emission is simpler. The scalar RTE is

$$\frac{dI(x, \mu)}{dx} = -I(x, \mu) + (1 - \omega_n) B_n(x). \tag{4.20}$$

Here, x is now the vertical optical thickness coordinate measured from the bottom of the layer, and $B_n(x)$ the Planck Function for temperatures in layer n. For the upwelling field, the solution corresponding to Eq. (4.13) is

$$\mathbf{I}(D_n, \mu) = \mathbf{I}(0, \mu)e^{-\rho_n D_n} + (1 - \omega_n) \int_0^{D_n} B_n(x)e^{-\rho_n(D_n-s)}ds. \tag{4.21}$$

There are two choices. The first is the plane-parallel option, with no sphericity along the line-of-sight. Once again, $x = -\rho_n z; s = (h_n - z)/\mu$. Using the linear parameterization $B_n(x) = a_n + b_n s$ in terms of the distance coordinate s, then Eq. (4.21) is then easy to integrate explicitly:

$$\int_0^{D_n} B_n(s)e^{-\rho_n(D_n-s)}ds = \frac{(\rho_n a_n - b_n)\left(1 - e^{-\rho_n D_n}\right)}{\rho_n^2} + \frac{b_n D_n}{\rho_n}. \tag{4.22}$$

In the second choice, the line-of-sight moves through a curved atmosphere, and we must again use the distance quadrature as in Eq. (4.19) to perform the integral in Eq. (4.21).

Building the field recursively using Eq. (4.15) also applies in the thermal case, except that the surface term is now given by $I_{Surf}^{\uparrow}(\Omega) = B_{surf} E(\mu)$, with emissivity $E(\mu)$ and surface Planck function B_{surf}. Downwelling treatment for direct thermal radiation unfolds in an analogous way.

Linearizations. We note that the FO fields are differentiable with respect to any basic set of optical properties that characterize the RT problem. In layer n, we assume that the IOPs $\{\Delta_n, \omega_n, \mathbf{\Pi}_n(\Omega, \Omega_0)\}$ have known analytic derivatives $\{\dot{\Delta}_n, \dot{\omega}_n, \dot{\mathbf{\Pi}}_n(\Omega, \Omega_0)\}$ with respect to variable ξ_n in layer n. Differentiation of Eq. (4.13) yields:

$$\dot{\mathbf{I}}(D_n, \Omega) = \left[\dot{\mathbf{I}}(0, \Omega) - \dot{\rho}_n D_n \mathbf{I}(0, \Omega)\right]e^{-\rho_n D_n} + \dot{\mathbf{S}}_n^{\uparrow}; \tag{4.23a}$$

$$\dot{\mathbf{S}}_n^{\uparrow} = \frac{1}{4\pi}\left[\dot{\sigma}_n \mathbf{\Pi}_n^{\uparrow}(\Omega, \Omega_0) + \sigma_n \dot{\mathbf{\Pi}}_n^{\uparrow}(\Omega, \Omega_0)\right]\mathbf{F}_{\odot} \int_0^{D_n} A_n(s)e^{-\rho_n(D_n-s)}ds$$

$$+ \frac{\sigma_n}{4\pi}\mathbf{\Pi}_n^{\uparrow}(\Omega, \Omega_0)\mathbf{F}_{\odot} \int_0^{D_n} \left[\dot{A}_n(s) - \dot{\rho}_n(D_n - s)A_n(s)\right]e^{-\rho_n(D_n-s)}ds. \tag{4.23b}$$

In terms of linearized IOPs, we have $\dot{\rho}_n = \dot{\Delta}_n / h_n$, and $\dot{\sigma}_n = \dot{\rho}_n \omega_n + \rho_n \dot{\omega}_n$. Equation (4.23b) contains some new integrals for the attenuations; the derivative attenuation from Eq. (4.14) is:

$$\dot{A}_n(s) = -d_{nn}(s)\dot{\rho}_n A_n(s). \tag{4.24}$$

Additionally, there are cross-layer derivatives of Eq. (4.14), since the attenuation depends on optical properties in layers $k < n$. If quantity ξ_k in layer k gives rise to optical property derivatives $\partial \rho_k / \partial \xi_k$, then (using a "prime" symbol for $\partial / \partial \xi_k$)

$$\mathbf{I}'(D_n, \Omega) = \mathbf{I}'(0, \Omega)e^{-\rho_n D_n} + \frac{\sigma_n}{4\pi} \mathbf{\Pi}_n^{\uparrow}(\Omega, \Omega_0)\mathbf{F}_{\odot} \int_0^{D_n} A'_{nk}(s)e^{-\rho_n(D_n - s)} ds. \tag{4.25}$$

Here, $A'_{nk}(s) = -d_{nk}(s)\rho'_k A_n(s)$, for $k < n$. Integrals with derivatives of the attenuations will have analytical expressions for the plane-parallel option and the pseudo-spherical option with the average-secant parameterization. These integrals have already cropped up in the full RTE linearizations. For the other option with full sphericity, integrals in Eq. (4.25) are done using the distance quadrature as noted above.

Solution linearization of the complete single-scatter field then follows from the chain-rule differentiation of the recurrence relation Eq. (4.15) and others similar to it. We note in particular the surface attenuation in Eq. (4.16) will have derivatives from every layer.

One final consideration is the surface-property linearization of the direct-reflection in Eq. (4.16). If the BRDF has derivative $\partial / \partial W$ (indicated below with a "double prime" symbol), with respect to some surface quantity W (e.g. wind speed for ocean glitter), then there will be Jacobians such as the following:

$$\mathbf{I}_0^{\uparrow ''}(\Omega) = C_N \mathbf{I}_{Surf}^{\uparrow ''}(\Omega); \quad \mathbf{I}_{Surf}^{\uparrow ''}(\Omega) = \frac{1}{\pi} A_{Surf} \mu_0 \mathbf{R}''(\Omega, \Omega_0)\mathbf{F}_{\odot}. \tag{4.26}$$

Linearization of the direct thermal FO field has two aspects. The first is differentiation with respect to the usual set of optical property variables—in a manner not unrelated to the above treatment for the SS field, we may differentiate the results in Eqs. (4.20)–(4.22), noting the need for additional quadrature integration involving the Plank functions. The other consideration is differentiation with respect to the Planck functions themselves—an important factor when considering temperature Jacobians in the thermal regime. This issue was dealt with in Sect. 3.3 and the facility developed there is appropriate to the present case.

4.5 F-Matrix and Phase Function Supplements

We deal with the vector F-Matrix (FMATRIX) case; the scalar supplement is a subset of this, as the phase function is the (1,1) element of the 4×4 scattering matrix. We suppose that there exists an external set of 6 F-matrix components $F_{11}(\Theta_k)$, $F_{12}(\Theta_k)$, $F_{22}(\Theta_k)$, $F_{33}(\Theta_k)$, $F_{34}(\Theta_k)$, $F_{44}(\Theta_k)$ on a grid of scattering angles $\{\Theta_k\}$. The first requirement of the "FMATRIX" supplement is to provide VLIDORT with F-matrix elements for the FO calculations, for a set of scattering angles $\{\Theta_v^{\pm}\}$ corresponding to choices of solar and viewing geometry configurations (indexed by the letter v). The second requirement is to generate sets of "Greek matrix" expansion coefficients to be used in the VLIDORT multiple scatter calculations.

For the first requirement, we use interpolation (currently, cubic splining is the default). It is important that the external grid $\{\Theta_k\}$ is fine enough in the forward scattering direction to allow for sufficient accuracy in the interpolation. For example, the tri-axial ellipsoidal dust–aerosol data set provided by Meng et al. (2010) has resolutions $0.01°$ from $0°$ to $2°$, $0.05°$ from $2°$ to $5°$, $0.1°$ from $5°$ to $10°$, $0.5°$ from $10°$ to $15°$, $1.0°$ from $15°$ to $175°$, and $0.25°$ for the remaining values to $180°$. Further, this data is presented as $\{\ln F_{11}(\Theta_k), R_{12}(\Theta_k), R_{22}(\Theta_k), R_{33}(\Theta_k), R_{34}(\Theta_k), R_{44}(\Theta_k)\}$, in terms of the logarithm of the phase function, and the ratios $R_{pq}(\Theta_k) = F_{pq}(\Theta_k)/F_{11}(\Theta_k)$. At the forward and backward peaks, the spline interpolation end-point gradients are constrained to take the values obtained by finite difference estimates using the data-set entries—we have found that this constraint makes a big improvement in interpolation accuracy, especially close to the forward peak.

The Greek coefficients are established through the orthogonality of the generalized spherical functions used for the Fourier-azimuth expansion of the phase matrix. Indeed, application of orthogonality properties gives (Mishchenko et al. 2006):

$$\beta_l = \left(l + \frac{1}{2}\right) \int_{-1}^{1} a_1(\mu) P_{00}^l(\mu) d\mu; \quad \zeta_l = \left(l + \frac{1}{2}\right) \int_{-1}^{1} a_4(\mu) P_{00}^l(\mu) d\mu; \quad \text{(4.27a)}$$

$$\alpha_l \pm \delta_l = \left(l + \frac{1}{2}\right) \int_{-1}^{1} [a_2(\mu) \pm a_3(\mu)] P_{2,\pm 2}^l(\mu) d\mu; \quad \text{(4.27b)}$$

$$\gamma_l = \left(l + \frac{1}{2}\right) \int_{-1}^{1} b_1(\mu) P_{02}^l(\mu) d\mu; \quad \epsilon_l = \left(l + \frac{1}{2}\right) \int_{-1}^{1} b_2(\mu) P_{02}^l(\mu) d\mu. \quad \text{(4.27c)}$$

Here $\mu = \cos \Theta$. The integrals are performed by double Gauss-Legendre quadrature with a high number of abscissae (typically 2000)—necessary to capture variability of the scattering functions especially near the forward peak. Clearly, a second round of interpolation is necessary to obtain scattering functions on the quadrature grid. This is another reason why it is highly desirable to have the input F-matrix functions at a sufficiently fine resolution—Greek coefficients will not be accurate otherwise.

5 Software Considerations

5.1 Package Overview

Here we give an overview of the current LIDORT code package organization (Fig. 1). The VLIDORT package is organized in a similar fashion, with one difference (there are two test directories). From the parent LIDORT directory, there are 9 upper-level subdirectories, including one for the main source code ("lidort_main"), one for LIDORT's Fortran 90 input/output type structure definition files ("lidort_def"), one for testing environments ("lidort_test"), one for the first-order (FO) source code ("fo_main_1p5") and one for the LIDORT supplement modules ("sup"). Further, object files ("obj"), module files ("mod"), LIDORT package utilities ("util"), and documentation ("docs") are also stored in separate directories.

In the main "test" directory, there are a number of driver programs which are designed to test most aspects of the code; these tests rely on configuration-file inputs to set the type-structure inputs. Details may be found in the User Guide. Test execution is controlled through a number of shell scripts. Test results should then be compared with pre-calculated results in the "saved_results" subdirectories.

The similarly-structured VLIDORT package has two test directories ("vlidort_s_test" and "vlidort_v_test"), the first for scalar-only output (no polarization), the second for vector output with polarization. Tests in the "vlidort_s_test" mirror those in LIDORT's "lidort_test" sub-directory.

As noted in the Introduction, earlier versions of the codes as described in [R1] were written in Fortran 77, and in transitioning to Fortran 90, all F77 COMMON
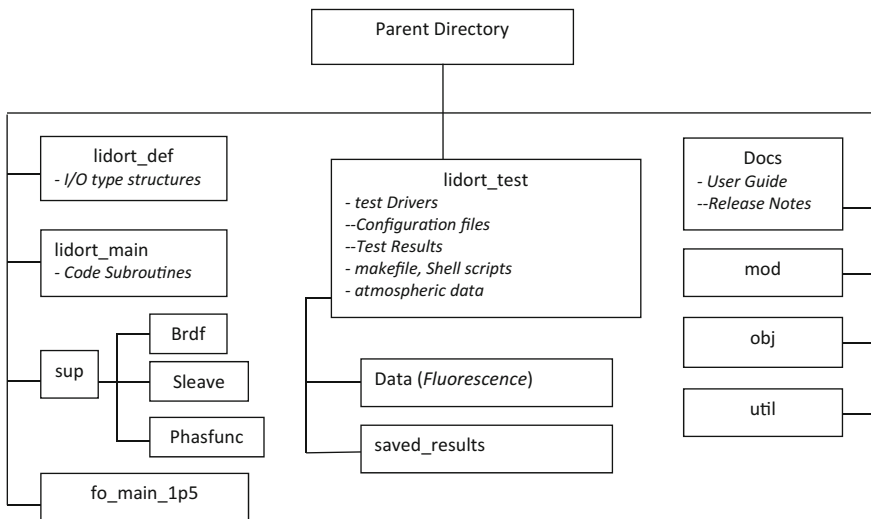
Fig. 1 Directory structure for the LIDORT package

blocks and associated "include" files were replaced. Now, all inputs and outputs are declared with the "implicit none" qualifier and the Fortran 90 "intent(in/out)" statements. Most recent versions of (V)LIDORT have the capability to run in the OpenMP parallel-computing environment, and we discuss this performance aspect in Sect. 5.2.3 below.

In addition, a new exception handling system has been introduced. Formerly, input-check and calculation errors were written to file as they occurred during model execution. This is not convenient for applications where (V)LIDORT is called within a larger software environment. Now, the (V)LIDORT models will assemble error and warning messages during execution, then output these messages along with "traces" to locate error sources.

5.2 Performance Enhancements

5.2.1 Multiple-SZA and "Observational Geometry" Options

In R1, the "multiple-SZA" facility was described. This takes advantage of the fact that the homogeneous RTE solution (in the absence of sources) needs to be done just once, whereas the particular integrals are required for every solar source beam. Earlier versions of (V)LIDORT were based on just one solar zenith angle (SZA) input, requiring the models to be called repeatedly for each desired beam geometry. Now, there is an internal loop over the set of solar sources. The time-saving arises when solving the boundary value problem $\mathbb{A}\mathbb{X}_v = \mathbb{B}_v$, where $\mathbb{X}_v$ is the vector of integration constants for the RTE solution with solar geometry v, $\mathbb{B}_v$ is the source term vector constructed from particular integral solutions for this geometry, with matrix $\mathbb{A}$ constructed only from RTE homogeneous solutions. The inverse matrix $\mathbb{A}^{-1}$ is calculated once only before the loop over solar geometry. Finding $\mathbb{A}^{-1}$ is the most time-consuming part of the code (especially for VLIDORT), and once done, the back-substitution exercise $\mathbb{X}_v = \mathbb{A}^{-1}\mathbb{B}_v$ is straightforward and fast. We remark only that convergence of the Fourier azimuth series in VLIDORT/LIDORT does depend on the SZA, so the model keeps track of this convergence separately for each SZA, thereby avoiding unwanted additional computation.

In the latest (V)LIDORT codes, we have introduced the "observational geometry" option. This arises from typical satellite atmosphere/surface retrieval algorithms, where there is a requirement for radiative transfer output at specific "solar zenith angle, viewing angle, relative azimuth angle" *observational geometry triplets*. Although (V)LIDORT models have multiple user-angle and multiple-SZA options, this capability is not efficient for generating output for observational geometry triplets. For example, if there are 4 such triplets, then previously (V)LIDORT was configured to generate $4 \times 4 \times 4 = 64$ output radiances, that is, one RT output for each of the 4 solar zenith angles, each of the 4 viewing angles, and each of the 4 relative azimuth angles. One may view this as computing a "$4 \times 4 \times 4$ lattice cube of solutions", and this is fine for building a look-up table (LUT). However for triplet

output with 4 SZA values, we require only those solutions along the diagonal of this "lattice cube of solutions" (i.e. 4 instead of 64, one for each triplet); the other 60 solutions are redundant.

The *observational geometry* facility bypasses this redundancy. When in force, a single call will generate the discrete-ordinate radiation fields for each SZA in the given triplet set, and then carry out post-processing only for those viewing zenith and relative azimuth angles uniquely associated with the triplet SZA. One of the big time savings here is with the internal geometry routines in (V)LIDORT—in our example, we require 4 calls instead of 64 calls needed in a "Lattice" computation for the same set of geometries.

5.2.2 Solution-Saving and BVP-Telescoping

"Solution-Saving" refers to the avoidance of finding unnecessary full solutions to the RTE when there is no scattering in a given layer n, for a given Fourier component m. Such a situation is quite common—for example in a Rayleigh (molecular) scattering atmosphere with one or more contiguous cloud layers. Rayleigh scattering has a $\cos^2 \Theta$ dependence on scattering angle Θ and is present only for Fourier $m \leq 2$, so in this example it is necessary to solve the RTE for $m > 2$ only in the cloud layers. For the non-scattering layers, the extinction law applies, and radiance is propagated through such layers according to transmittances $T_n(\mu) = \exp[-\Delta_n/\mu]$, where Δ_n is the layer vertical optical depth, and μ any directional cosine. Linearized RTE solutions are also zero in non-scattering layers, but any linearized solutions in adjacent scattering layers will also be transmittance-propagated through non-scattering layers. Note that if this transmittance propagation passes through layer n for which a parameter derivative $\dot{\Delta}_n = \partial \Delta_n / \partial \xi_n$ is present, then the linearization will pick up an additional term $\dot{T}_n(\mu) = -T_n(\mu)\dot{\Delta}_n/\mu$.

The "BVP-Telescoping" facility was first described in [R1]. Here we give a brief summary, noting additional developments in the latest versions. BVP-Telescoping goes hand-in-hand with "Solution-Saving"—when there is a contiguous group of actively scattering layers in an otherwise non-scattering atmosphere (e.g. Rayleigh for Fourier $m > 2$), it is possible to reduce the size of the boundary value problem to a linear algebra system of much lower rank—essentially we are determining discrete-ordinate RTE integration constants for the radiation fields in these active scattering layers, noting that these solutions will be transmittance-propagated through non-scattering layers above and below the cloud deck. This reduction of rank in the BVP linear-algebra system represents a substantial time-saving operation. Performance aspects of the telescoped BVP *ansatz* have been examined also in (Efremenko et al. 2013).

For Fourier components $m > 0$, a Lambertian (isotropic) lower boundary surface is dark, and in this case light propagated downwards from the cloud deck will not be reflected upwards from the surface to contribute to the upwelling field in the cloud. This is the simplest situation for BVP telescoping, and this has been the default until recently. Now, the (V)LIDORT codes are able to deal with a reflecting surface

characterized by a BRDF. In this case downwardly-propagated scattered light from the cloud deck will be reflected upwards to rejoin the active layers, and one must fold in a BRDF surface-reflection condition for the lowest boundary of the reduced BVP. Linearization of this more general BVP scenario with surface reflection has also been accomplished. Note that the BVP telescoping will also work with a cloud deck having no non-scattering layers beneath ("fog at the ground" scenario).

5.2.3 OpenMP Facility, Software Speed-Up Improvements

The VLIDORT 2.7 (2015) and LIDORT 3.7 (2014) codes were given the capability to run in OpenMP distributed parallel-computing environments. OpenMP compatibility requires that the codes are "thread-safe", and this development has necessitated some software structural changes. The main such change has been the removal of "SAVE" statements—arrays with this attribute would be shared among parallel OpenMP threads and this would lead to anomalous behavior.

Care should be taken regarding proper preparation for OpenMP test environments: parallel programming tests such as these can result in memory segmentation faults if steps are not taken to ensure enough memory is set aside for both main and OpenMP-spawned computational threads. For more information, the reader should consult the VLIDORT and LIDORT User Guides, where there are some guidelines for using the codes in OpenMP.

Both codes show excellent "scalability" when using OpenMP. Running with 2 cores in a single CPU, the speed-up is very close to 2.0, while an OpenMP run with 4 cores will generate a speed-up of around 3.8. These figures are typical, but will depend on the application (large wavelength loops in hyperspectral environments, for example).

Finally, there have been a number of internal coding revisions to ensure that do-loop ordering and array-filling makes efficient use of memory. This latter consideration has proved to be most valuable for applications with profile Jacobians, where performances have improved by as much as 30%.

5.3 Benchmarking, Validation and Inter-model Consistency

5.3.1 Rayleigh and Aerosol Slab Problems

Benchmarking for the VLIDORT model was discussed in [R1], and we give a brief summary here. The *Rayleigh slab problem* refers to a single-layer plane-parallel medium of optical thickness 1.0, conservative Rayleigh scattering with polarization (single scattering albedo $\omega = 1.0$), and a Lambertian lower surface. For this scenario, the V-component of the Stokes vector is decoupled. The original solution for linear polarization was found by Chandrasekhar in the 1940s, and benchmark results were presented in the Coulson, Dave and Sekera (CDS) tables (Coulson et al. 1960).

Although VLIDORT is technically unable to handle conservative scattering mathematically, it has been possible to validate VLIDORT by setting $\omega = 0.99999999$, and as noted in [R1], VLIDORT can reproduce all CDS entries to within the accuracy level specified for these tables.

More recently, a new set of Rayleigh slab-problem benchmarks has been generated (Natraj and Hovenier 2012), using a new coding of Chandrasekhar's classical Rayleigh slab solution. The coding uses double-precision arithmetic, and these authors showed that the CDS tables are generally accurate to only 4 or 5 significant figures (SF). These authors also carried out a VLIDORT validation for the CDS inputs; using 32 discrete ordinates in the half-space, and they were able to get agreement to better than 9 SF with VLIDORT output for this problem.

The (plane-parallel) *aerosol slab problem* is similar, but with aerosol scattering instead of Rayleigh. For VLIDORT, we first compared with results from the discrete-ordinate model of Siewert (2000b). This scenario has a solar angle of 53.130° ($\mu_0 = 0.6$), with single scatter albedo $\omega = 0.973527$, surface albedo 0.0, slab optical thickness 1.0, and a set of Greek coefficients as noted in Table 1 of Siewert (2000b). Stokes 4-vector output is given in a set of eight tables for a number of optical thickness values from 0 to 1, and a number of output streams. Multiple scattering was based on 24 discrete ordinates in the half space. All results were reproduced by VLIDORT with complete fidelity. A second benchmarking was done against results from Garcia and Siewert (1989) for a related aerosol slab problem; this time with albedo 0.1. Tables 3–10 in Garcia and Siewert (1989) were reproduced to 5 SF in all cases, with VLIDORT using 20 discrete ordinates in the half space. This result is noteworthy because the radiative transfer computations in Garcia and Siewert (1989) were done using a completely different radiative transfer methodology (the so-called F_N method).

5.3.2 LIDORT-VLIDORT Comparisons

The two models should be completely consistent in all respects. Indeed, with the exception of additional polarized optical properties in VLIDORT, the model inputs are the same. VLIDORT will operate in "scalar mode" when the input parameter NSTOKES is set to 1 and in this case, VLIDORT reproduces very closely any results generated independently from calls to LIDORT with equivalent geometrical and optical inputs. It should be noted that exact agreement (to *all* SF) cannot be expected, since LIDORT uses Green's function methods to solve the RTE particular integral, whereas VLIDORT employs the conventional exponential-substitution method for this task. We have found that agreement between VLIDORT and LIDORT in all situations is better than 9 SF for this scalar-only comparison.

5.3.3 Consistency with the 2-Stream and Linearized Rotational-Raman Scattering Models

LIDORT and VLIDORT are the two main models in the family of RT codes provided by RT Solutions. One of the other codes of interest here is the dedicated 2-stream (2S) radiance/flux model (Spurr and Natraj 2011). 2S is a multiple-scattering RT model in a multi-layer stratified media; it can operate with either choice of the 2-stream ordinate ($\bar{\mu} = 0.5$ or $\bar{\mu} = 1/\sqrt{3}$). 2S has most of the capabilities of the LIDORT/VLIDORT codes, although flux and radiance outputs are presently confined to level boundaries. 2S has both BRDF and surface-leaving supplements which are fully compatible with those of LIDORT.

In the context of the accelerated radiative transfer applications using principal component analysis (PCA) of optical property fields (Natraj et al. 2010) (see also Spurr et al. 2016 and references therein), both VLIDORT and LIDORT have been independently validated against 2S, for top-of-atmosphere upwelling radiances and bottom of atmosphere downwelling fields. This was achieved by setting the number of discrete ordinates $N_d = 1$ in LIDORT and VLIDORT, and comparing the multiple scattering output with 2S output for a given solar zenith angle and a set of viewing scenarios.

The second model to check for consistency is the LRRS (LIDORT Rotational-Raman Scattering) code (Spurr et al. 2008). This is a scalar LIDORT-based RT model with both elastic scattering (no wavelength redistribution) and inelastic (wavelength redistributed) molecular scattering due to rotational-Raman transitions by air molecules. LRRS treats elastic scattering to all orders, but Raman photons are inelastically scattered just once. LRRS is able to ingest the same optical inputs as LIDORT, so the elastic scattering output from LRRS can be compared exactly with its counterpart from LIDORT. This is true for a wide range of applications, since LRRS has both BRDF and Surface-leaving supplements which are fully compatible with LIDORT. LRRS has recently been linearized and brought up to date with the latest developments in (V)LIDORT, and the LRRS upgrades will be described in a monograph under preparation (Spurr and Christi 2018).

6 Concluding Remarks

In this paper, we have reviewed developments and upgrades to the LIDORT and VLIDORT discrete ordinate radiative transfer models over the last 10 years, since the publication of the last review paper [R1] (Spurr 2008). We have also included a summary of the newer FO (First-Order) codes, which are stand-alone but fully compatible with the (V)LIDORT models.

We have focused in particular on a number of supplements, including modules for the generation of BRDF inputs, solar-induced fluorescence, a new treatment of water-leaving radiances, and external single-scattering computations based on user-defined phase functions and scattering matrices (the latter in connection with the FO model).

We also discussed some new linearization capabilities, especially the generation of analytic weighting functions with respect to Planck function inputs as required for thermal emission sources, and the derivation of level-boundary Jacobians (with respect to temperatures and mixing ratios in particular) from transformations of the (V)LIDORT layer optical property Jacobians.

We have summarized a number of performance enhancements, including a generalized Taylor-series treatment for closely adjacent polar-stream directions, improvements to the treatment of BRDFs in BVP-Telescoping scenarios. Both the LIDORT and VLIDORT models are now "thread-safe", which enables them to be used in OpenMP parallel computing environments.

The latest release versions for the codes are: Version 2.8 for VLIDORT, Version 3.8 for LIDORT, Version 2.4 for 2S, Version 1.5 for the FO code, and Version 2.5 for the LRRS code. These versions have the same capabilities, with a completely consistent set of inputs and outputs. The LIDORT family of RT models has found widespread use in the remote sensing community—the codes are freely available and covered by the GNU General GPL license version 3.0. All codes are accompanied by User Guide documentation, and may be obtained by contacting Robert Spurr at RT Solutions Inc.

Appendix A. RTE Solutions for VLIDORT

A.1 Azimuthal Separation

In applications involving randomly oriented particles with one plane of symmetry, the scattering matrix $\mathbf{F}(\Theta)$ in Eq. (2.5) has six independent entries:

$$
\mathbf{F}(\Theta) = \begin{pmatrix} a_1(\Theta) & b_1(\Theta) & 0 & 0 \\ b_1(\Theta) & a_2(\Theta) & 0 & 0 \\ 0 & 0 & a_3(\Theta) & b_2(\Theta) \\ 0 & 0 & -b_2(\Theta) & a_4(\Theta) \end{pmatrix}.
\tag{A.1}
$$

For this form of the scattering matrix, one can develop expansions of these functions in terms of a set of generalized spherical functions $P_{mn}^l(\cos\Theta)$ (Mishchenko et al. 2006):

$$a_1(\Theta) = \sum_{l=0}^{M} \beta_l P_{00}^l(\cos\Theta);$$

$$a_2(\Theta) + a_3(\Theta) = \sum_{l=0}^{M} (\alpha_l + \zeta_l) P_{2,2}^l(\cos\Theta);$$

$$a_2(\Theta) - a_3(\Theta) = \sum_{l=0}^{M} (\alpha_l - \zeta_l) P_{2,-2}^l(\cos\Theta);$$

$$a_4(\Theta) = \sum_{l=0}^{M} \delta_l P_{00}^l(\cos\Theta) \tag{A.2}$$

$$b_1(\Theta) = \sum_{l=0}^{M} \gamma_l P_{02}^l(\cos\Theta); \quad b_2(\Theta) = -\sum_{l=0}^{M} \epsilon_l P_{02}^l(\cos\Theta) \tag{A.3}$$

The (1,1) entry in $\mathbf{F}(\Theta)$ is just the phase function and satisfies the normalization condition:

$$\frac{1}{2}\int_0^\pi a_1(\Theta)\sin\Theta d\Theta = 1. \tag{A.4}$$

The set of six expansion coefficients ("Greek constants") $\{\alpha_l, \beta_l, \gamma_l, \delta_l, \varepsilon_l, \zeta_l\}$ are key inputs to VLIDORT, and the number of terms M in these expansions depends on the desired level of numerical accuracy. Here, $\{\beta_l\}$ are the phase function expansion coefficients as used in the scalar RTE. These "Greek constants" specify the polarized-light single-scattering law, and there are a number of efficient analytical techniques for their computation, not only for spherical particles (see for example de Rooij and van der Stap 1984) but also for randomly-oriented homogeneous and inhomogeneous non-spherical particles and aggregated scatterers (Hovenier et al. 2004; Mackowski and Mishchenko 1996; Mishchenko and Travis 1998).

To proceed with the RTE solution, it is necessary to make Fourier decompositions (in terms of the cosine and sine of the relative azimuth angle between incident and scattered light directions) of the phase matrix and the Stokes vector in order to separate the azimuthal dependence.

$$\mathbf{I}(x,\mu,\phi-\phi') = \frac{1}{2}\sum_{m=0}^{M}(2-\delta_{m,0})\boldsymbol{\Phi}_m(\phi-\phi')\mathbf{I}_m(x,\mu); \tag{A.5}$$

$$\boldsymbol{\Pi}(\mu,\phi,\mu',\phi') = \frac{1}{2}\sum_{m=0}^{M}(2-\delta_{m,0})[\mathbf{C}_m(\mu,\mu')\cos m(\phi-\phi') + \mathbf{S}_m(\mu,\mu')\sin m(\phi-\phi')]; \tag{A.6}$$

Here, $\boldsymbol{\Phi}_m(\phi) \equiv Diag[\cos m\phi, \cos m\phi, \sin m\phi, \sin m\phi]$. We follow the formulation for azimuthal separation of the scattering matrix developed by Siewert (1982),

Vestrucci and Siewert (1984). Most vector radiative transfer models now follow this work. Accordingly, we write:

$$\mathbf{C}_m(\mu, \mu') = \mathbf{A}_m(\mu, \mu') + \mathbf{D}\mathbf{A}_m(\mu, \mu')\mathbf{D}; \ \mathbf{S}_m(\mu, \mu') = \mathbf{A}_m(\mu, \mu')\mathbf{D} - \mathbf{D}\mathbf{A}_m(\mu, \mu');$$
(A.7)

$$\mathbf{A}_m(\mu, \mu') = \sum_{l=m}^{M} \mathbf{P}_l^m(\mu)\mathbf{B}_l\mathbf{P}_l^m(\mu'); \ \mathbf{D} = Diag[1, 1, -1, -1];$$
(A.8)

$$\mathbf{B}_l = \begin{pmatrix} \beta_l & \gamma_l & 0 & 0 \\ \gamma_l & \alpha_l & 0 & 0 \\ 0 & 0 & \delta_l & -\varepsilon_l \\ 0 & 0 & \varepsilon_l & \zeta_l \end{pmatrix}; \mathbf{P}_l^m(\mu) = \begin{pmatrix} P_l^m(\mu) & \gamma_l & 0 & 0 \\ \gamma_l & R_l^m(\mu) & -T_l^m(\mu) & 0 \\ 0 & -T_l^m(\mu) & R_l^m(\mu) & 0 \\ 0 & 0 & 0 & P_l^m(\mu) \end{pmatrix}.$$
(A.9)

The "Greek matrices" contain the spherical-function expansion coefficients, while the matrices $\mathbf{P}_l^m(\mu)$ contain the associated Legendre functions $P_l^m(\mu)$ and the functions $R_l^m(\mu)$ and $T_l^m(\mu)$ which are closely related to the generalized spherical functions $P_{mn}^l(\mu)$ (for details, see for example Siewert 2000b).

This azimuth separation process yields the following RTE for the Fourier component $\mathbf{I}_m(x, \mu)$:

$$\mu\frac{d\mathbf{I}_m(x, \mu)}{dx} = -\mathbf{I}_m(x, \mu) + \frac{\omega}{2}\sum_{l=m}^{M}\mathbf{P}_l^m(\mu)\mathbf{B}_l\int_{-1}^{1}\mathbf{P}_l^m(\mu')\mathbf{I}_m(x, \mu')d\mu' + \mathbf{Q}_m(x, \mu).$$
(A.10)

For the solar source term, with solar direction $\{-\mu_0, \phi_0\}$, we have.

$$\mathbf{Q}_m^{\odot}(x, \mu) = \frac{\omega}{2}\sum_{l=m}^{M}\mathbf{P}_l^m(\mu)\mathbf{B}_l\mathbf{P}_l^m(-\mu_0)\mathbf{F}_{\odot}\exp[-\tau_{\odot}(x, \mu)],$$
(A.11)

in terms of the TOA solar flux $\mathbf{F}_{\odot} = [F_{\odot}, 0, 0, 0]^{\mathrm{T}}$ and solar beam attenuation $\exp[-\tau_{\odot}(x, \mu_0)]$, where $\tau_{\odot}(x, \mu_0)$ is the beam optical depth in a spherical-shell atmosphere.

A.2 Homogeneous Solutions in VLIDORT

A.2.1 Eigenvalue Solutions

In the discrete-ordinate solution method, we solve for each Fourier component in Eq. (A.10) by first finding the homogeneous solutions (without the solar or thermal

source terms). With the familiar discrete ordinate quadrature with stream directions and weights $\{\pm\mu_i, c_i\}$ for $i = 1, \ldots N_d$, where N_d is the number of discrete ordinates in the polar half space. The resulting vector RTE for these streams is then

$$\pm\mu_i \frac{d\mathbf{I}_m(x, \pm\mu_i)}{dx} \pm \mathbf{I}_m(x, \pm\mu_i) = \frac{\omega}{2} \sum_{l=m}^{M} \mathbf{P}_l^m(\pm\mu_i)\mathbf{B}_l \sum_{j=1}^{N_d} \left[\mathbf{P}_l^m(\mu_j)\mathbf{I}_m(x, \mu_j) \right.$$
$$\left. + \mathbf{P}_l^m(-\mu_j)\mathbf{I}_m(x, -\mu_j) \right] \tag{A.12}$$

There are $8N_d$ coupled first-order linear differential equations for $\mathbf{I}_m(x, \pm\mu_i)$ in this system, which is solved by eigenvalue methods, using the *ansatz*:

$$\mathbf{I}_\alpha(x, \pm\mu_i) = \mathbf{W}_\alpha(\pm\mu_i) \exp[-k_\alpha x]. \tag{A.13}$$

We then define the following two vectors of rank $4N_d$:

$$\mathbb{W}_\alpha^\pm = \left[\mathbf{W}_\alpha^{\mathrm{T}}(\pm\mu_1), \mathbf{W}_\alpha^{\mathrm{T}}(\pm\mu_2), \ldots, \mathbf{W}_\alpha^{\mathrm{T}}(\pm\mu_{N_d}) \right]^{\mathrm{T}}. \tag{A.14}$$

The system is decoupled using sum and difference vectors $\mathbb{X}_\alpha = \mathbb{W}_\alpha^+ + \mathbb{W}_\alpha^-$ and $\mathbb{Y}_\alpha = \mathbb{W}_\alpha^+ - \mathbb{W}_\alpha^-$, and the order is then be reduced from $8N_d$ to $4N_d$. The result is an eigenproblem for the collection of separation constants and associated solution vectors $\{k_\alpha, \mathbb{X}_\alpha\}$, where index $\alpha = 1, \ldots, 4N_d$ labels the eigensolutions. The eigenmatrix for this system is constructed from the optical property inputs $\{\omega, \mathbf{B}_l\}$ and combination products of matrices $\mathbf{P}_l^m(\pm\mu_i)$. The eigenproblem is Siewert (2000b):

$$\mathbb{X}_\alpha^*\mathbb{G} = k_\alpha^2\mathbb{X}_\alpha^*; \mathbb{G}\mathbb{X}_\alpha = k_\alpha^2\mathbb{X}_\alpha; \tag{A.15a}$$

$$\mathbb{G} = \mathbb{F}^+\mathbb{F}^-; \mathbb{F}^\pm = \left[\mathbb{E} - \frac{\omega}{2} \sum_{l=m}^{M} \mathbb{P}(l, m)\mathbf{B}_l\mathbf{A}^\pm\mathbb{P}^{\mathrm{T}}(l, m)\mathbb{C} \right]\mathbb{M}^{-1}; \tag{A.15b}$$

$$\mathbb{P}(l, m) = Diag\left[\mathbf{P}_l^m(\mu_1), \mathbf{P}_l^m(\mu_2), \ldots, \mathbf{P}_l^m(\mu_{N_d}) \right]^{\mathrm{T}}; \tag{A.15c}$$

$$\mathbb{M} = Diag\left[\mu_1\mathbf{E}, \mu_2\mathbf{E}, \ldots, \mu_{N_d}\mathbf{E} \right]; \mathbb{C} = Diag\left[c_1\mathbf{E}, c_2\mathbf{E}, \ldots, c_{N_d}\mathbf{E} \right]. \tag{A.15d}$$

Here, $\mathbf{A}^\pm = \mathbf{E} \pm (-1)^{l-m}\mathbf{D}; \mathbf{E} = Diag[1, 1, 1, 1]; \mathbf{D} = Diag[1, 1, -1, -1]$, $\mathbb{E}$ is the $4N_d \times 4N_d$ identity matrix; $\mathbb{X}_\alpha^*$ and $\mathbb{X}_\alpha$ are the left and right eigenvectors respectively, with $\mathbb{X}_\alpha^*$ the conjugate transpose of $\mathbb{X}_\alpha$. The link between $\mathbb{X}_\alpha$ and solution vectors $\mathbb{W}_\alpha^\pm$ comes through the auxiliary equations:

$$\mathbb{W}_\alpha^\pm = \frac{1}{2}\mathbb{M}^{-1}\left[\mathbb{E} \pm \frac{1}{k_\alpha}\mathbb{F}^+ \right]\mathbb{X}_\alpha. \tag{A.16}$$

Eigenvalues occur in pairs $\{\pm k_\alpha\}$. Left and right eigenvectors share the same spectrum of eigenvalues. As noted by Siewert (2000b), both complex- and real-variable eigensolutions may be present in the full Stokes 4-vector case (rank $4N_d$). Eigensolutions may be determined numerically with the complex-variable eigensolver DGEEV

from the LAPACK suite (Anderson et al. 1995). DGEEV generates both right and left eigenvectors, which have unit moduli. In the scalar case (no polarization, solutions only for the I-component of the Stokes vector), the eigensystem has rank N_d, the eigenmatrix is symmetric and all eigensolutions are real-valued. In this case, the eigensolver module ASYMTX (Stamnes et al. 1988) may be used. ASYMTX is an adaptation of the LAPACK routine for real-valued problems; it delivers only the right eigenvectors. For vector solutions neglecting circular polarization (Stokes I, Q and U only), complex eigensolutions are absent, and one may then use the faster ASYMTX routine.

Returning to the full Stokes 4-vector case, the complete homogeneous solution in one layer is then:

$$\mathbb{I}^+(x) = \mathbb{D}^+ \sum_{\alpha=1}^{4N_d} \left\{ \mathfrak{L}_\alpha \mathbb{W}_\alpha^+ \exp[-k_\alpha x] + \mathfrak{M}_\alpha \mathbb{W}_\alpha^- \exp[-k_\alpha(\Delta - x)] \right\}; \quad \text{(A.17a)}$$

$$\mathbb{I}^-(x) = \mathbb{D}^- \sum_{\alpha=1}^{4N_d} \left\{ \mathfrak{L}_\alpha \mathbb{W}_\alpha^- \exp[-k_\alpha x] + \mathfrak{M}_\alpha \mathbb{W}_\alpha^+ \exp[-k_\alpha(\Delta - x)] \right\}; \quad \text{(A.17b)}$$

Here, $\mathbb{D}^- = Diag[\mathbf{D}, \mathbf{D}, \ldots \mathbf{D}]$, and $\mathbb{D}^+ = \mathbb{E}$; these matrices arise from application of symmetry relations (Siewert 2000b). The use of optical thickness $\Delta - x$ in the second exponential ensures that solutions remain bounded (Stamnes and Conklin 1984). The quantities $\{\mathfrak{L}_\alpha, \mathfrak{M}_\alpha\}$ are the constants of integration; determined by application of the boundary conditions and solution of the resulting boundary-value problem.

In the Stokes 4-vector case, some contributions to $\mathbb{I}^\pm(x)$ will be complex, some real. It is understood that we compute the real parts of any complex variable expressions. Thus for example if $\{k_\alpha, \mathbb{W}_\alpha^-\}$ is a complex eigensolution with associated (complex) integration constant $\mathfrak{L}_\alpha$, the real part of the solution will be:

$$\text{Re}\left[\mathfrak{L}_\alpha \mathbb{W}_\alpha^- e^{-k_\alpha x}\right] = \text{Re}[\mathfrak{L}_\alpha]\text{Re}\left[\mathbb{W}_\alpha^- e^{-k_\alpha x}\right] - \text{Im}[\mathfrak{L}_\alpha]\text{Im}\left[\mathbb{W}_\alpha^- e^{-k_\alpha x}\right]. \quad \text{(A.18)}$$

From a bookkeeping standpoint, one must keep count of the number of real and complex solutions, and treat them separately in the numerical implementation. For clarity of exposition, we have not made an explicit separation of complex variables, and it will be clear from the context whether real or complex variables are under consideration.

A.2.2 Linearization of the Eigenvalue Solutions

For a single layer, we require derivatives of the eigensolution $\{k_\alpha, \mathbb{W}_\alpha^\pm\}$ with respect to some atmospheric variable ξ in that layer. The starting point for the differentiation is the set of linearized optical properties $\mathcal{V} \equiv \xi \partial\Delta/\partial\xi; \mathcal{U} \equiv \xi\partial\omega/\partial\xi; \mathcal{Z}_l \equiv \xi\partial\mathbf{B}_l/\partial\xi$, that is, normalized partial derivatives of the set of IOPs $\{\Delta, \omega, \mathbf{B}_l\}$. The

eigensolution depends only on the quantities $\{\omega, \mathbf{B}_l\}$. Applying the linearization operator $\mathcal{L} \equiv \xi \partial/\partial \xi$ to the eigenmatrix, we find

$$\mathcal{L}(\mathbb{G}) = \mathcal{L}(\mathbb{F}^+)\mathbb{F}^- + \mathbb{F}^+\mathcal{L}(\mathbb{F}^-); \tag{A.19a}$$

$$\mathcal{L}(\mathbb{F}^\pm) = -\left[\frac{1}{2}\{\mathcal{U}\mathbb{P}_{lm}\mathbf{B}_l + \omega\mathbb{P}_{lm}\mathcal{Z}_l\}\mathbf{A}^\pm\mathbb{P}_{lm}^{\mathrm{T}}\mathbb{C}\right]\mathbb{M}^{-1}. \tag{A.19b}$$

Linearization treatments are different for the full Stokes 4-vector case, and the scalar and Stokes 3-vector situations.

4-vector case. We apply linearization to both the left and right eigensystems:

$$\mathcal{L}(\mathbb{X}_\alpha^*)\mathbb{G} + \mathbb{X}_\alpha^*\mathcal{L}(\mathbb{G}) = 2k_\alpha\mathcal{L}(k_\alpha)\mathbb{X}_\alpha^* + k_\alpha^2\mathcal{L}(\mathbb{X}_\alpha^*); \tag{A.20a}$$

$$\mathbb{G}\mathcal{L}(\mathbb{X}_\alpha) + \mathcal{L}(\mathbb{G})\mathbb{X}_\alpha = 2k_\alpha\mathcal{L}(k_\alpha)\mathbb{X}_\alpha + k_\alpha^2\mathcal{L}(\mathbb{X}_\alpha). \tag{A.20b}$$

We form a dot product $\otimes$ by pre-multiplying the second of these equations by the transpose vector $\mathbb{X}_\alpha^*$:

$$\mathbb{X}_\alpha^* \otimes \mathbb{G}\mathcal{L}(\mathbb{X}_\alpha) + \mathbb{X}_\alpha^* \otimes \mathcal{L}(\mathbb{G})\mathbb{X}_\alpha = 2k_\alpha\mathcal{L}(k_\alpha)\mathbb{X}_\alpha^* \otimes \mathbb{X}_\alpha + k_\alpha^2\mathbb{X}_\alpha^* \otimes \mathcal{L}(\mathbb{X}_\alpha). \tag{A.21}$$

Using the relation $\mathbb{X}_\alpha^* \otimes \mathbb{G}\mathcal{L}(\mathbb{X}_\alpha) = \mathbb{X}_\alpha^*\mathbb{G} \otimes \mathcal{L}(\mathbb{X}_\alpha) = k_\alpha^2\mathbb{X}_\alpha^* \otimes \mathcal{L}(\mathbb{X}_\alpha)$, we find that

$$y_\alpha \equiv \mathcal{L}(k_\alpha) = \frac{\mathbb{X}_\alpha^* \otimes \mathcal{L}(\mathbb{G})\mathbb{X}_\alpha}{2k_\alpha\mathbb{X}_\alpha^* \otimes \mathbb{X}_\alpha}. \tag{A.22}$$

This is the linearization of the separation constants. Next, we substitute this result in Eq. (A.20a) to obtain the following linear algebra problem for each eigensolution linearization:

$$\mathbb{H}_\alpha\mathcal{L}(\mathbb{X}_\alpha) = \mathbb{C}_\alpha; \quad \mathbb{H}_\alpha = \mathbb{G} - k_\alpha^2\mathbb{E}; \quad \mathbb{C}_\alpha = 2k_\alpha y_\alpha\mathbb{X}_\alpha - \mathcal{L}(\mathbb{G})\mathbb{X}_\alpha. \tag{A.23}$$

For real eigensolutions, this system has rank $4N_d$, and for complex solutions, the rank is $8N_d$.

Implementation of this system of equations "as is" is not possible due to the degeneracy of the eigenproblem, and we need additional constraints to find the unique solution for $\mathcal{L}(\mathbb{X}_\alpha)$. The treatment for real and complex solutions is different.

For the real-valued eigensolutions, the unit-modulus eigenvector normalization is $\mathbb{X}_\alpha \otimes \mathbb{X}_\alpha = 1$ in dot-product notation. Linearizing, this yields one equation:

$$\mathcal{L}(\mathbb{X}_\alpha) \otimes \mathbb{X}_\alpha + \mathbb{X}_\alpha \otimes \mathcal{L}(\mathbb{X}_\alpha) = 0. \tag{A.24}$$

The solution procedure uses $4N_d - 1$ equations from Eq. (A.23), along with Eq. (A.24) to form a slightly modified linear system of rank $4N_d$. This system is then

solved by standard means using the DGETRF and DGETRS LU-decomposition routines from the LAPACK suite.

For complex-valued eigensolutions, Eq. (A.23) is a complex-variable system for both the real and imaginary parts of the linearized eigenvectors. There are $8N_d$ equations in all, but now we require two constraint conditions to remove the eigenproblem arbitrariness. The first is Eq. (24). The second condition is imposed by the following normalization from the LAPACK DGEEV routine for solution of complex-valued eigenproblems: for that element of a (complex-valued) eigenvector which has the largest real value, the corresponding imaginary part is always set to zero. Thus, for eigenvector $\mathbb{X}_\alpha$ with components $X_j \in \mathbb{X}_\alpha$, $j = 1, 2, \ldots 4N_d$, if X_q satisfies condition $\mathrm{Re}[X_q] = \max_{j=1,2,\ldots 4N_d}\{\mathrm{Re}[X_j]\}$, then $\mathrm{Im}[X_q] = 0$, and consequently $\mathcal{L}(\mathrm{Im}[X_q]) = 0$. This is the second condition. The solution procedure is then: (1) in Eq. (A.24) to strike out the qth row and column in matrix $\mathbb{H}_\alpha$ for which $\mathrm{Im}[X_q]$ is zero, and the qth column in vector $\mathbb{C}_\alpha$; and (2) in the resulting system of rank $8N_d - 1$, replace one of the rows with the normalization constraint Eq. (A.24). $\mathcal{L}(\mathbb{X}_\alpha)$ is then the solution of the resulting linear system.

Scalar and 3-vector case. Here the (real-valued) eigensolutions are obtained using eigensolver ASYMTX—this has no adjoint solution, so there is no determination of $\mathcal{L}(k_\alpha)$ as in Eq. (A.22). Instead, we solve for variables $\{\mathcal{L}(k_\alpha), \mathcal{L}(\mathbb{X}_\alpha)\}$ using $\mathbb{H}_\alpha \mathcal{L}(\mathbb{X}_\alpha) = 2k_\alpha \mathcal{L}(k_\alpha)\mathbb{X}_\alpha - \mathcal{L}(\mathbb{G})\mathbb{X}_\alpha$ from above, plus the normalization condition to form a joint system of rank $3N_d + 1$ (vector) or rank $N_d + 1$ (scalar).

Having derived the linearizations $\{\mathcal{L}(k_\alpha), \mathcal{L}(\mathbb{X}_\alpha)\}$, we complete this section by differentiating the auxiliary result in Eq. (A.16) to establish $\mathcal{L}(\mathbb{W}_\alpha^\pm)$:

$$\mathcal{L}(\mathbb{W}_\alpha^\pm) = \frac{1}{2}\mathbb{M}^{-1}\left[\mp\frac{\mathcal{L}(k_\alpha)}{k_\alpha^2}\mathbb{F}^+ \pm \frac{1}{k_\alpha}\mathcal{L}(\mathbb{F}^+)\right]\mathbb{X}_\alpha + \frac{1}{2}\mathbb{M}^{-1}\left[\mathbb{E} \pm \frac{1}{k_\alpha}\mathbb{F}^+\right]\mathcal{L}(\mathbb{X}_\alpha).$$
(A.25)

Finally, we have linearizations of the transmittance derivatives in Eq. (A.17a):

$$\mathcal{L}(\exp[-k_\alpha x]) = -[x\mathcal{L}(k_\alpha) + k_\alpha \mathcal{L}(x)]\exp[-k_\alpha x].$$
(A.26)

Since the partial layer optical thickness x is proportional to the total layer optical depth Δ in an optically uniform layer, we have $\mathcal{L}(x) = x/\Delta \mathcal{L}(\Delta) = x/\Delta \mathcal{V}$ in terms of the basic linearized optical property input $\mathcal{V} \equiv \xi \partial \Delta / \partial \xi$.

A.3 Particular Integral Solar Solutions in VLIDORT

In VLIDORT, particular integrals for both solar and thermal sources in the vector RTE are established using traditional substitution methods, rather than the Green's function approach which is used in LIDORT. This is mainly for bookkeeping reasons

associated with the use of complex and real variables. In this section we discuss solutions for the solar sources using this method.

A.3.1 Solar-Beam Sources—Substitution Method

For the solar source term with direction $\{-\mu_0, \phi_0\}$, we have from Eq. (A.11) above, the following source terms in the discrete ordinate directions:

$$\mathbf{Q}_{nm}^{\odot}(x, \pm\mu_i) = \frac{\omega_n}{2} \sum_{l=m}^{M} \mathbf{P}_l^m(\pm\mu_i)\mathbf{B}_{nl}\mathbf{P}_l^m(-\mu_0)\mathbf{F}_{\odot}\mathrm{T}_n\exp[-\lambda_n x]. \qquad (A.27)$$

Here, we have kept the layer index explicit, and x denotes the vertical optical thickness as measured from the top of layer n; we used the pseudo-spherical treatment of solar beam attenuation (Sect. 3.4). The exponential form for the beam attenuation allows us to write the particular solution in the form:

$$\mathbf{I}_{nm}^{\odot}(x, \pm\mu_i) = \mathbf{Z}_n(\pm\mu_i)\mathrm{T}_n\exp[-\lambda_n x], \qquad (A.28)$$

and by analogy with the homogeneous case, we may define the following vectors of rank $4N_d$:

$$\mathbb{Z}_n^{\pm} = \left[\mathbf{Z}_n^{\mathrm{T}}(\pm\mu_1), \mathbf{Z}_n^{\mathrm{T}}(\pm\mu_2), \ldots, \mathbf{Z}_n^{\mathrm{T}}(\pm\mu_{N_d})\right]^{\mathrm{T}}. \qquad (A.29)$$

We decouple the equations using sum and difference vectors $\mathbb{R}_n = \mathbb{Z}_n^+ + \mathbb{Z}_n^-$ and $\mathbb{S}_n = \mathbb{Z}_n^+ - \mathbb{Z}_n^-$, and the order is reduced from $8N_d$ to $4N_d$. We obtain the following linear-algebra problem of rank $4N_d$:

$$\mathbb{H}_n\mathbb{R}_n = \mathbb{B}_n; \quad \mathbb{H}_n = \lambda_n^2\mathbb{E} - \mathbb{G}_n; \quad \mathbb{B}_n = \left[\mathbb{F}_n^-\mathbb{Q}_n^+ + \frac{1}{\lambda_n}\mathbb{Q}_n^-\right]\mathbb{M}^{-1}; \qquad (A.30)$$

$$\mathbb{Q}_n^{\pm} = \omega_n \sum_{l=m}^{M} \mathbb{P}_{\odot}(l, m)\mathbf{B}_{nl}\mathbf{A}^{\pm}\mathbb{P}_{\odot}^{\mathrm{T}}(l, m)\mathbb{M}^{-1}; \qquad (A.31a)$$

$$\mathbb{P}_{\odot}(l, m) = Diag\left[\mathbf{P}_l^m(-\mu_0), \mathbf{P}_l^m(-\mu_0), \ldots, \mathbf{P}_l^m(-\mu_0)\right]^{\mathrm{T}}. \qquad (A.31b)$$

Matrices $\mathbb{F}_n^-$ and $\mathbb{G}_n$ were defined in Eqs. (A.15).

This system is solved using the LU-decomposition modules DGETRF and DGETRS from LAPACK; the formal solution is $\mathbb{X}_n^{\odot} = \mathbb{H}_n^{-1}\mathbb{B}_n$. The particular integral is completed with the auxiliary equations:

$$\mathbb{Z}_n^{\pm} = \frac{1}{2}\mathbb{M}^{-1}\left[\mathbb{E} \pm \frac{1}{\lambda_n}\mathbb{F}_n^+\right]\mathbb{R}_n. \qquad (A.32)$$

In the scalar LIDORT model, this system has rank N_d. In the vector model, the particular solution consists only of real variables.

A.3.2 Linearized Solar-Beam Sources—Substitution Method

For this linearization, the most important consideration is the presence of cross-derivatives: in a fully illuminated atmosphere, the particular solution in layer n is differentiable with respect to atmospheric variables ξ_p in all layers $p \leq n$. For the solar beam attenuation, transmittance T_n depends on variables ξ_p in layers $p < n$, while average secant λ_n depends on variables ξ_p in layers $p \leq n$. Linearization of the average-secant parameterization is in Sect. 3.4. It follows that the solution vectors $\mathbb{Z}_n^{\pm}$ will also depend on ξ_p for $p \leq n$, so *their* linearizations will contain cross-derivatives. Finally, we note that the eigenmatrix $\mathbb{G}_n$ is constructed from optical properties only defined in layer n, so that $\mathcal{L}_p(\mathbb{G}_n) = 0, \forall p \neq n$.

Differentiation of Eqs. (A.30) and A.31a) yields a related linear problem:

$$\mathbb{H}_n \mathcal{L}_p(\mathbb{R}_n) = \mathbb{B}'_{np} = \mathcal{L}_p(\mathbb{B}_n) - \mathcal{L}_p(\mathbb{H}_n)\mathbb{R}_n; \tag{A.33a}$$

$$\mathcal{L}_p(\mathbb{H}_n) = -\delta_{np}\mathcal{L}_p(\mathbb{G}_n) + 2\lambda_n \mathcal{L}_p(\lambda_n)\mathbb{E}; \tag{A.33b}$$

$$\mathcal{L}_p(\mathbb{B}_n) = \delta_{np}\left[\mathcal{L}_n(\mathbb{F}_n^-)\mathbb{Q}_n^+ + \mathbb{F}_n^- \mathcal{L}_n(\mathbb{Q}_n^+) + \frac{1}{\lambda_n}\mathcal{L}_n(\mathbb{Q}_n^-)\right]\mathbb{M}^{-1} - \frac{\mathcal{L}_p(\lambda_n)}{\lambda_n^2}\mathbb{Q}_n^-\mathbb{M}^{-1}; \tag{A.33c}$$

$$\mathcal{L}_n(\mathbb{Q}_n^{\pm}) = \sum_{l=m}^{M}\left[\mathcal{U}_n\mathbb{P}_{\odot}(l,m)\mathbf{B}_{nl} + \omega_n\mathbb{P}_{\odot}(l,m)\boldsymbol{\mathcal{Z}}_{nl}\right]\mathbf{A}^{\pm}\mathbb{P}_{\odot}^{\mathrm{T}}(l,m)\mathbb{M}^{-1}. \tag{A.33d}$$

In Eq. (A.33c), the quantity $\mathcal{L}_n(\mathbb{F}_n^-)$ comes from Eq. (A.19b). This linear system has the same matrix $\mathbb{H}_n$, but with a different source vector $\mathbb{B}'_{np}$. The solution is then found by back-substitution, given that the inverse of $\mathbb{H}_n$ has already been established when solving for $\mathbb{R}_n$. Thus, $\mathcal{L}_p(\mathbb{R}_n) = \mathbb{H}_n^{-1}\mathbb{B}'_{np}$. Linearization of the particular integral is then completed through differentiation of the auxiliary equations (A.32):

$$\mathcal{L}_p(\mathbb{Z}_n^{\pm}) = \frac{1}{2}\mathbb{M}^{-1}\left[\mathbb{E} \pm \frac{1}{\lambda_n}\mathbb{F}_n^+\right]\mathcal{L}_p(\mathbb{R}_n) \pm \frac{1}{2\lambda_n^2}\mathbb{M}^{-1}\left[\lambda_n\delta_{np}\mathcal{L}_n(\mathbb{F}_n^+) - \mathcal{L}_p(\lambda_n)\mathbb{F}_n^+\right]\mathbb{R}_n. \tag{A.34}$$

Appendix B. BRDF Kernel Functions

B.1 Ocean Glitter Kernels

B.1.1 Cox-Munk Glint Reflectance

For ocean glitter, we use the well-known geometric-optics regime for a single rough-surface redistribution of incident light, in which the reflection function is governed by Fresnel reflectance and takes the form (Jin et al. 2006):

$$\rho_{CM}\big(\mu, \mu', \varphi - \varphi', m, \sigma^2\big) = \frac{F(m, \theta_r)}{\mu\mu'|\gamma_r|^4} P\big(\gamma_r, \sigma^2\big) D\big(\mu, \mu', \sigma^2\big) \tag{B.1}$$

Here, σ^2 is the slope-squared variance (also known as the mean-slope-square) of the Gaussian probability function $P\big(\gamma_r, \sigma^2\big)$ which has argument γ_r (the polar direction of the reflected beam), $D\big(\mu, \mu', \sigma^2\big)$ is a shadow function correction (Sancer 1969).

$F(m, \theta_r)$ is the scalar Fresnel reflection for incident angle $\theta_r = \frac{1}{2}\theta_{SCAT}$ and relative refractive index m. The scattering angle θ_{SCAT} is determined in the usual manner from the incident and reflected directional cosines μ' and μ, and the relative azimuth $\varphi - \varphi'$.

The two non-linear parameters characterizing this kernel are $\big\{m, \sigma^2\big\}$. The probability and shadow functions are given by:

$$P\big(\alpha, \sigma^2\big) = \frac{1}{\pi\sigma^2}\exp\left[-\frac{\alpha^2}{\sigma^2\big(1 - \alpha^2\big)}\right]. \quad D\big(\alpha, \beta, \sigma^2\big) = \frac{1}{1 + \Lambda\big(\alpha, \sigma^2\big) + \Lambda\big(\beta, \sigma^2\big)};$$

$$\Lambda\big(\alpha, \sigma^2\big) = \frac{1}{2}\left\{\sqrt{\frac{\big(1 - \alpha^2\big)}{\pi}}\frac{\sigma}{\alpha}\exp\left[-\frac{\alpha^2}{\sigma^2\big(1 - \alpha^2\big)}\right] - erfc\left[-\frac{\alpha}{\sigma\sqrt{\big(1 - \alpha^2\big)}}\right]\right\}.$$
$$\tag{B.2}$$

The variance is commonly related to the wind speed W in (m/s) through the empirical relation $\sigma^2 = 0.003 + 0.00512W$ (Cox and Munk 1954a). A typical value for m is 1.33.

For the linearization, the key parameter is the wind-speed (or equivalently, the mean slope-square) The probability function is easily differentiated with respect to σ^2. Indeed, we have:

$$\frac{\partial P\big(\alpha, \sigma^2\big)}{\partial\sigma^2} = \frac{P\big(\alpha, \sigma^2\big)}{\sigma^4}\left[\frac{\alpha^2}{\big(1 - \alpha^2\big)} - \sigma^2\right]. \tag{B.3}$$

Again, the shadow function can be differentiated analytically with respect to σ^2 in a straightforward manner, once we recall that the derivative of the error function is Gaussian. We thus have linearized BRDF quantities prepared for (V)LIDORT, which will then be able to generate analytic weighting functions with respect to the wind speed.

Linearization of the kernel with respect to refractive index m will require the partial derivative $\partial F(m, \theta_r)/\partial m$, which is easy to determine from the usual Fresnel formula; this Jacobian is less useful.

We note also that VLIDORT has a vector kernel function for sea-surface glitter reflectance, based on the specification in Mishchenko and Travis (1998).

This formulation is for a single Fresnel reflectance by wave facets. In reality, glitter is the result of many reflectances. It is possible to incorporate a correction for

multiple reflectances for this glitter contribution, both for the direct-bounce term and the Fourier components. Given that the glitter maximum is typically dominated by direct reflectance of the solar beam, we confine discussion to multiple-reflectance of this term. We consider only one extra order of reflectance:

$$R(\Omega, \Omega_0) \cong R_0(\Omega, \Omega_0) + R_1(\Omega, \Omega_0); R_1(\Omega, \Omega_0) = \int\limits_0^{2\pi} \int\limits_0^1 R_0(\Omega, \Omega') R_0(\Omega', \Omega_0) d\mu' d\varphi'.$$

(B.4)

$R_0(\Omega, \Omega')$ is the zeroth-order reflectance from incident direction $\Omega' = \{\mu', \varphi'\}$ to reflected direction $\Omega = \{\mu, \varphi\}$. The azimuthal integration is done by double Gaussian quadrature over the intervals $[-\pi, 0]$ and $[0, \pi]$; the polar stream integration is also done by quadrature. It is obviously possible to calculate higher orders of reflectance for all BRDF functions, but this rapidly becomes computationally prohibitive. We have found that the neglect of multiple glitter reflectances can lead to errors of 1–3% in the upwelling intensity at the top of the atmosphere, the higher figures being for larger solar zenith angles. Finally, we note that the higher-order reflectances are in turn differentiable with respect to the slope-squared parameter, so that Jacobians for the wind speed can still be determined.

B.1.2 Alternative Cox-Munk Glint Reflectance

The above ocean-glint reflectance option in LIDORT is based on an older implementation of the well-known Cox-Munk distribution of wave facets—this does not include any treatment of whitecaps (foam), and there is no allowance for the wind direction. Further, the CM implementation is based on a fixed real-valued refractive index for water.

We have now added an alternative Cox-Munk implementation which addresses these issues. This new-CM option is based on the glint treatment in the 6S code (Vermote et al. 1997; Kotchenova et al. 2006), and includes an empirical whitecap contribution. The latter model also includes a more recent treatment of water-leaving radiance (Morel and Gentili 2009), and the LIDORT SLEAVE supplement has been updated according to the 6S treatment. These new 6S-based options in the LIDORT BRDF and SLEAVE supplements are designed to operate in tandem. Indeed, the total surface radiance in the 6S model is given by

$$I(\mu_1, \mu_0, \varphi_1 - \varphi_0) = (1 - R_{wc})S(\mu_1, \mu_0) + R_{wc} + (1 - R_L)\rho_{NCM}(\mu_1, \mu_0, \varphi_1 - \varphi_0), \quad \text{(B.5)}$$

where the water-leaving term $S(\mu_1, \mu_0)$ does depend on the outgoing and incoming directions but is azimuth-independent, ρ_{NCM} is the Cox-Munk glint reflectance, and R_{wc} and R_L are the empirically-derived whitecap contributions (final and Lambertian).

For LIDORT to possess this functionality for the ocean surface, the BRDF supplement must provide the second and third terms on the right-hand-side of (B.5), while the SLEAVE supplement supplies the first term. Note that both supplements

require the same whitecap formulation. The derivation of the water-leaving term is done in the next section.

The Cox-Munk calculation for ρ_{NCM} depends on the wind speed, wind-direction and refractive index. The latter is a complex variable that depends on the salinity of the ocean, and we use the 6S formulation. The anisotropic treatment assumes an (azimuthal) wind direction relative to the solar incident beam, and this follows the formulae developed by Cox and Munk in the 1950s.

For facet anisotropy, the wind-direction is dependent on the solar position, and it follows that any BRDF quantity will pick up additional dependence on the solar angle. It is then only possible to use this "New-CM" option with a single solar zenith angle (and hence a single wind-direction azimuth) - multiple solar beams are ruled out. See also the remark in Sect. 4.2.3 regarding use of the black-sky albedo. There is exception handling for this eventuality.

In addition, we have linearized this "New-CM" glint reflectance with respect to the wind speed. Finally, we note that this glint reflectance is scalar only, so that the above considerations apply only to the (1,1) element of the reflectance matrix in the vector BRDF supplement for VLIDORT. Polarization of this contribution is currently undergoing investigation.

B.2 Land-Surface BRDF Kernels

B.2.1 MODIS BRDF System

The five MODIS-type kernels (numbers 2–6 in Table 3) (Wanner et al. 1995) must be used in a linear combination with a Lambertian kernel. The kernels divide naturally into two groups: the *volume scattering* terms with no non-linear parameters (Ross-thin, Ross-thick) and the *geometric-optics* terms with 2 non-linear parameters (Li-sparse, Li-dense) or no non-linear parameters (Roujean). See Wanner et al. (1995) and Spurr (2004) for details of the kernel formulae.

In fact, it is standard practice in MODIS BRDF retrievals to use a combination of Lambertian, Ross-thick and Li-Sparse kernels, and this 3-kernel combination is common:

$$\rho_{total}\big(\Omega, \Omega'\big) = f_{iso} + f_{vol}\rho_{RossThick}\big(\Omega, \Omega'\big) + f_{geo}\rho_{LiSparse}\big(\Omega, \Omega'\big) \qquad \text{(B.6)}$$

An alternative form of the Ross kernels has also been introduced—this has a better parameterization of the hot-spot effect. The Rahman and Hapke kernels (#7 or #8) were discussed in [R1].

B.2.2 BPDF Kernels

The 3 BPDF kernels (Maignan et al. 2009) (numbers 10–12 in Table 3) were developed as semi-empirical models for polarized land-surface bidirectional reflectances. All three kernels are based on Fresnel reflectance. The polarization effects enter through the Fresnel reflectance term; for LIDORT, we require only the scalar Fresnel reflectance.

For the BPDF "Soil" type, the scalar reflectance is:

$$\rho_{SOIL}(\mu, \mu', \varphi - \varphi', m) = \frac{F(m, \theta_r)}{4\mu\mu'}(1 - \sin\theta_r). \tag{B.7}$$

As before, $F(m, \theta_r)$ is the Fresnel reflection for incident angle $\theta_r = \frac{1}{2}\theta_{SCAT}$ and relative refractive index m (which is the sole kernel parameter). Linearization of the kernel with respect to refractive index m requires the partial derivative of $F(m, \theta_r)$.

For the BDPF "Vegetation" type, there is dependence on the leaf orientation probability $\sigma(\alpha)$ and the leaf facet projections, through use of a plagiophile distribution:

$$\rho_{VEGN}(\mu, \mu', \varphi - \varphi', m) = \frac{F(m, \theta_r)}{4\mu\mu'}\frac{\zeta(\alpha)}{H}(1 - \sin\theta_r);$$

$$\cos\alpha = \frac{(\mu + \mu')}{2\cos\theta_r}; \quad \zeta(\alpha) = \frac{16}{\pi}\cos^2\alpha\,\sin\alpha;$$

$$H = \frac{\sum_{k=0}^{3} a_k\mu^k}{\mu} + \frac{\sum_{k=0}^{3} a_k\mu'^k}{\mu'}, \tag{B.8}$$

where the leaf projection H depends on the set of "plagiophile coefficients" $\{a_k\}$. Again, the refractive index is the only surviving kernel parameter to be considered for linearization.

For the BPDF "NDVI" kernel, we have:

$$\rho_{NDVI}(\mu, \mu', \varphi - \varphi', m, N, C) = \frac{CF(m, \theta_r)}{4(\mu + \mu')}\exp[-\tan\theta_r]\exp[-N], \tag{B.9}$$

where there are exponential attenuation terms, one of which depends on the NDVI N; in this formula, the scaling factor is C (nominally, this is set to 1.0). Linearizations with respect to the parameters N and C are easy to establish. The NDVI varies from -1 to 1 and is defined as the ratio of the difference to the sum of two radiance measurements, one in the visible and one in the infrared.

Appendix C. Taylor Series Expansions

We have already noted that certain "multipliers" arising from optical-thickness integrations which are needed to find various solutions to components of the RT field

may possess instability when certain limiting conditions are in place. Taylor series expansions are then required to remove such instabilities. Although some simple expansions were used in earlier versions of the LIDORT and VLIDORT codes, the whole issue has been completely revised in recent versions of the code (since [R1] was published), and we go into some detail here. Remarks on the VLIDORT implementation are made where appropriate.

C.1 Multipliers for the Intensity Field

We first look at the homogeneous-solution and primary-scatter *downwelling* multipliers, and the Green's function (downwelling) multiplier for the diffuse source term at discrete-ordinate polar cosines. These are respectively:

$$H_j^\downarrow = \frac{\left(e^{-\Delta k_j} - e^{-\Delta \mu^{-1}}\right)}{\mu^{-1} - k_j}; F^\downarrow = \frac{\left(e^{-\Delta \mu^{-1}} - e^{-\Delta \lambda}\right)}{\lambda - \mu^{-1}}; G_j^\downarrow = \frac{\left(e^{-\Delta k_j} - e^{-\Delta \lambda}\right)}{\lambda - k_j}. \tag{C.1}$$

As before, λ is the layer average secant corresponding to spherical-shell attenuation of the solar beam, k_j is the separation constant corresponding to discrete ordinate stream j arising from solution of the homogeneous RTE eigenproblem, μ is the cosine of the line-of-sight viewing polar angle, and Δ is the layer total optical depth. [Layer index n is suppressed for clarity].

For polarized RT with VLIDORT, some of the separation constants k_j may be complex variables; but λ and μ are always real-valued. Taylor series expansions involving k_j are only applicable for real values.

We consider the limiting cases $|\lambda - k_j| \to 0$, $|\mu^{-1} - k_j| \to 0$ or $|\lambda - \mu^{-1}| \to 0$. Writing ϵ for any of these quantities, if the order of the Taylor-series expansion is M, then we neglect terms $O\left(\epsilon^{M+1}\right)$. Considering first the multiplier $G_j^\downarrow$ from Eq. (C.1), then if $\epsilon = \lambda - k_j$, we have $e^{-\Delta k_j} = We^{\epsilon \Delta} \approx W\mathbf{z}_{M+1}(\Delta)$, where we have written $W = e^{-\Delta \lambda}$, and we have used the notation $\mathbf{z}_{M+1}(x) = \sum_{m=0}^{M+1} z_m(x)\epsilon^m$ to approximate the exponential $e^{\epsilon \Delta}$, so that the coefficients are $z_0(x) = 1, z_m(x) = x^m/m!$ for $0 < m \le M + 1$. Applying the expansion, we find:

$$G_j^\downarrow \approx \frac{W(\mathbf{z}_{M+1}(\Delta) - 1)}{\epsilon} = \Delta W\mathbf{z}_M^*(\Delta). \tag{C.2}$$

Here, the new coefficients are $z_0^* = 1, z_m^* = \Delta^m/(m + 1)!$ for $0 < m \le M$. Note that we need to expand first to order $M + 1$ to ensure that the final expression is defined to order M.

Next we look at those two multipliers for the Green's function post-processed field which may be susceptible to instability through appearance of the term $\lambda - k_j$ in the denominator:

$$M_j^{\uparrow\uparrow} = \frac{H_j^{\uparrow} - F^{\uparrow}}{\lambda - k_j}; \quad M_j^{\downarrow\downarrow} = \frac{H_j^{\downarrow} - F^{\downarrow}}{\lambda - k_j}, \tag{C.3}$$

where $H_j^{\downarrow}$ and $F^{\downarrow}$ have been defined in Eq. (1), and

$$H_j^{\uparrow} = \frac{\left(1 - e^{-\Delta k_j} e^{-\Delta \mu^{-1}}\right)}{\mu^{-1} + k_j}; F^{\uparrow} = \frac{\left(1 - e^{-\Delta \mu^{-1}} e^{-\Delta \lambda}\right)}{\lambda + \mu^{-1}}. \tag{C.4}$$

Looking at $M_j^{\uparrow\uparrow}$ in Eq. (C.3), we set $k_j = \lambda - \epsilon$, and find:

$$M_j^{\uparrow\uparrow} \approx \frac{Y}{\epsilon}[(1 - WU\mathbf{z}_{M+1}(\Delta))\mathbf{a}_{M+1}(Y) - (1 - WU)], \tag{C.5}$$

where $U = e^{-\Delta \mu^{-1}}$, $Y = \left(\lambda + \mu^{-1}\right)^{-1}$, and series $\mathbf{a}_{M+1}(x)$ has coefficients $a_0(x) = 1, a_m(x) = x^m$ for $0 < m \leq M + 1$. Since $a_0 = z_0 = 1$, it is apparent that the $(1 - WU)$ contributions will fall out, and the result can then be written:

$$M_j^{\uparrow\uparrow} \approx Y[Y\mathbf{a}_M(Y) - WU\mathbf{c}_M(\Delta, Y)]. \tag{C.6}$$

Here $\mathbf{c}_M(\Delta, Y)$ has coefficients $c_0 = 1$, and $c_m(\Delta, Y) = \sum_{p=0}^{m} z_p(\Delta)a_{m-p}(Y)$ for $0 < m \leq M$. We have found that use of these series coefficients is convenient for computation, allowing us to generate expressions to any order of accuracy without complicated algebraic expressions.

The other multiplier $M_j^{\downarrow\downarrow}$ in Eq. (C.3) may be treated similarly. Note that multipliers in Eq. (C.4) are numerically stable entities, along with the other Green's function multipliers $G_j^{\uparrow}$, $M_j^{\uparrow\uparrow}$ and $M_j^{\uparrow\downarrow}$, the latter three quantities being defined with denominator $\lambda + k_j$.

The above multipliers are required for output of the upwelling and downwelling radiance fields at layer boundaries. LIDORT has the ability to generate output at any level away from layer boundaries (the "partial-layer" option). In this case, source function integration to an optical thickness $\tau < \Delta$ in layer n will result in "partial-layer" multipliers similar to those already defined above. For example, consider the following three functions:

$$H_j^{\downarrow}(\tau) = \frac{\left(e^{-\tau k_j} - e^{-\tau \mu^{-1}}\right)}{\mu^{-1} - k_j}; F^{\downarrow}(\tau) = \frac{\left(e^{-\tau \mu^{-1}} - e^{-\tau \lambda}\right)}{\lambda - \mu^{-1}}; G_j^{\downarrow}(\tau) = \frac{\left(e^{-\tau k_j} - e^{-\tau \lambda}\right)}{\lambda - k_j}; \tag{C.7}$$

These are the downwelling partial-layer multipliers equivalent to those in Eq. (C.1). Taylor series expansions for these and the corresponding Green's function multipliers have been generated in a similar fashion.

C.2 Linearized Multipliers for the Jacobian Fields

The entire LIDORT discrete ordinate solution is analytically differentiable (Spurr 2002) with respect to any atmospheric quantity, and this includes the multipliers discussed above. It is therefore necessary to develop Taylor series expansions for those *linearized* multipliers which are susceptible to numerical instability.

We start with partial derivatives $\dot{k}_j \equiv \partial k_j/\partial \xi$, $\dot{\lambda} \equiv \partial \lambda/\partial \xi$ and $\dot{\Delta} \equiv \partial \Delta/\partial \xi$ with respect to some atmospheric quantity ξ defined in layer n (index is not explicit here). Since the layer optical depth Δ is an intrinsic optical property, its derivative $\dot{\Delta}$ is also an intrinsic input. The viewing angle cosine μ has no derivative. Note that $\dot{k}_j = 0$ and $\dot{\Delta} = 0$ for a *profile* atmospheric quantity ξ_m defined in layer $m \neq n$ (no cross-layer derivatives); on the other hand, the average secant λ will have cross-layer derivatives for layers $m < n$, thanks to solar beam attenuation through the atmosphere. For the third multiplier in Eq. (C.1), the derivative is

$$\frac{\partial G_j^{\downarrow}}{\partial \xi} = \frac{-e^{-\Delta k_j}\left(\dot{k}_j\Delta + k_j\dot{\Delta}\right) + e^{-\Delta\lambda}\left(\dot{\lambda}\Delta + \lambda\dot{\Delta}\right) - G_j^{\downarrow}\left(\dot{\lambda} - \dot{k}_j\right)}{\lambda - k_j}. \qquad \text{(C.8)}$$

Expanding Eq. (C.8) as a Taylor series with $k_j = \lambda - \epsilon$, and using the series-coefficient notation developed in Sect. C.1, we find:

$$\frac{\partial G_j^{\downarrow}}{\partial \xi} \approx \frac{-\left(\dot{k}_j\Delta + \lambda\dot{\Delta} + \epsilon\dot{\Delta}\right)W\mathbf{z}_{M+1} + \left(\dot{\lambda}\Delta + \lambda\dot{\Delta}\right)W - \left(\dot{\lambda} - \dot{k}_j\right)\Delta W\mathbf{z}_{M+1}^*}{\epsilon}. \qquad \text{(C.9)}$$

Again, $W = e^{-\Delta\lambda}$, and the series $\mathbf{z}_{M+1}$ and $\mathbf{z}_{M+1}^*$ have argument Δ and were defined in Sect. C.1. Since $z_0 = z_0^* = 1$, the lowest-order terms in the numerator of Eq. (C.9) will fall out, and we are left with:

$$\frac{\partial G_j^{\downarrow}}{\partial \xi} \approx -W\left[\dot{\Delta}\mathbf{z}_M + \left(\dot{k}_j\Delta + \lambda\dot{\Delta}\right)\Delta\mathbf{z}_M^* + \left(\dot{\lambda} - \dot{k}_j\right)\Delta^2\mathbf{z}_M^{\dagger}\right], \qquad \text{(C.10)}$$

where the third series $\mathbf{z}_M^{\dagger}$ has coefficients $z_0^{\dagger} = 1/2$ and $z_m^{\dagger} = \Delta^m/(m+2)!$ for $0 < m \leq M$. Note that the presence of the series $\mathbf{z}_M^{\dagger}$ to order M implies that the original series must be computed to order $M + 2$; that is, we require $\mathbf{z}_{M+2}$.

Linearization of the Green's function multipliers in Eq. (C.3) follows similar considerations. We give one example; explicit differentiation of $M_j^{\uparrow\uparrow}$ yields

$$\frac{\partial M_j^{\uparrow\uparrow}}{\partial \xi} = \frac{\dot{H}_j^{\uparrow} - \dot{F}^{\uparrow} - \left(\dot{\lambda} - \dot{k}_j\right)M_j^{\uparrow\uparrow}}{\lambda - k_j}; \qquad \text{(C.11)}$$

$$\dot{H}_j^\uparrow = \frac{Ue^{-\Delta k_j}\big[\dot{\Delta}(\mu^{-1}+k_j)+\dot{k}_j\Delta\big]-\dot{k}_jH_j^\uparrow}{\mu^{-1}+k_j};\dot{F}^\uparrow = \frac{UW\big[\dot{\Delta}(\mu^{-1}+\lambda)+\dot{\lambda}\Delta\big]-\dot{\lambda}F^\uparrow}{\lambda+\mu^{-1}}.$$
$$(C.12)$$

Expanding in the usual manner and employing results established already, we find

$$\dot{H}_j^\uparrow \approx \big[UW(\dot{k}_j\Delta+\dot{\Delta}Y^{-1}-\epsilon\dot{\Delta})\mathbf{z}_N - \dot{k}_jY(1-WU\mathbf{z}_N)\mathbf{a}_N\big]Y\mathbf{a}_N; \qquad (C.13)$$
$$\dot{F}^\uparrow = \big[UW(\dot{\lambda}\Delta+\dot{\Delta}Y^{-1})-\dot{\lambda}Y(1-WU)\big]Y; \qquad (C.14)$$

where now $N = M + 2$ terms in the series have been retained in the expansions. Eq. (C.11) implies that we also require the original multiplier $M_j^{\uparrow\uparrow}$ expanded to order $M + 1$, that is,

$$M_j^{\uparrow\uparrow} \approx Y[Y\mathbf{a}_{M+1}(Y)-WU\mathbf{c}_{M+1}(\Delta,Y)] \qquad (C.15)$$

should be used in Eq. (C.11). Putting together the above three equations, and canceling out the lowest-order terms in the expansions, we find after some algebra that

$$\frac{\partial M_j^{\uparrow\uparrow}}{\partial \xi} \approx Y\Big[WU\mathbf{S}_M^{(1)}-\mathbf{S}_M^{(2)}-\mathbf{S}_M^{(3)}\Big], \qquad (C.16)$$

where the three series $\mathbf{S}_M^{(q)} = \sum_{m=0}^M s_m^{(q)}\epsilon^m$ have coefficients

$$s_m^{(1)} = \big(\dot{\Delta}Y^{-1}+\dot{k}_j\Delta\big)c_{m+1}(\Delta,Y)-\dot{\Delta}c_m(\Delta,Y); \qquad (C.17a)$$
$$s_m^{(2)} = Y\dot{k}_j[b_{m+1}(Y)+WUd_{m+1}(\Delta,Y)]; \qquad (C.17b)$$
$$s_m^{(3)} = \big(\dot{\lambda}-\dot{k}_j\big)[b_{m+2}(Y)+WUc_{m+2}(\Delta,Y)]. \qquad (C.17c)$$

Subsidiary coefficients for $0 < m \le M$ are given by $a_m = Y^m$ (series expansion of $(1-\epsilon Y)^{-1}$), and $b_m = (m+1)Y^m$ (series expansion of $(1-\epsilon Y)^{-2}$). We also have the product coefficients $c_m(\Delta,Y) = \sum_{p=0}^m z_p(\Delta)a_{m-p}(Y)$ and $d_m(\Delta,Y) = \sum_{p=0}^m z_p(\Delta)b_{m-p}(Y)$ obtained from the first exponential series $\mathbf{z}(\Delta)$ which approximates $e^{\Delta\epsilon}$.

Derivatives of the other multipliers subject to possible instability may be obtained similarly, and there are also derivatives of the partial-layer multipliers to be considered.

Software for all these instability cases has been written for LIDORT and VLIDORT - in both models, the order M controls the accuracy of the Taylor series expansions. The other parameter controlling the use of these limiting-case calculations is the "small-number" value ϵ. LIDORT and VLIDORT use double-precision floating-point arithmetic. With this in mind, we have chosen $\varepsilon = 10^{-3}$ as the default, after testing linearized multiplier accuracies obtained by running the model with and without the instability corrections. In practice $M = 3$ provides more than sufficient accuracy for this choice of ϵ.

References

Anderson E, Bai Z, Bischof C, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Ostrouchov A, Sorensen D (1995) LAPACK User's Guide, 2nd edn. Philadephia, Society for Industrial and Applied Mathematics

Chandrasekhar S (1960) Radiative transfer. Dover Publications Inc., New York

Chami M, Santer R, Dilligeard E (2001) Radiative transfer model for the computation of radiance and polarization in an ocean-atmosphere system: polarization properties of suspended matter for remote sensing. Appl Opt 40:2398–2416

Coulson K, Dave J, Sekera D (1960) Tables related to radiation emerging from planetary atmosphere with Rayleigh scattering. University of California Press, Berkeley

Cox C, Munk W (1954a) Statistics of the sea surface derived from sun glitter. J Mar Res 13:198–227

Cox C, Munk W (1954b) Measurement of the roughness of the sea surface from photographs of the sun's glitter. J Opt Soc Am 44:838–850

de Rooij WA, Van der Stap CCAH (1984) Expansion of Mie scattering matrices in generalized spherical functions, Astron Astrophys 131:237–248

Doicu A, Trautmann T (2009) Two linearization methods for atmospheric remote sensing. J Quant Spectrosc Radiat Transf 110:477–490

Efremenko D, Doicu A, Loyola D, Trautmann T (2013) Acceleration techniques for the discrete ordinate method. J Quant Spectrosc Radiat Transf 114:73–81

Frankenberg C, O'Dell C, Guanter L, McDuffie J (2012) Remote sensing of near-infrared chlorophyll fluorescence from space in scattering atmospheres: implications for its retrieval and interferences with atmospheric CO_2 retrievals. Atmos Meas Tech 5:2081–2094

Garcia RDM, Siewert CE (1989) The F_N method for radiative transfer models that include polarization. J Quant Spectrosc Radiat Transf 41:117–145

Hapke B (1993) Theory of reflectance and emittance spectroscopy. Cambridge University Press, Cambridge, UK

Hasekamp OP, Landgraf J (2002) A linearized vector radiative transfer model for atmospheric trace gas retrieval. J Quant Spectrosc Radiat Transf 75:221–238

Hovenier JW, Van der Mee C, Domke H (2004) Transfer of polarized light in planetary atmospheres: Basic concepts and practical methods, Kluwer, Dordrecht

Huang H, Qin W, Spurr RJ, Liu Q (2017) Evaluation of atmospheric effects on land surface directional reflectance with the coupled RAPID and VLIDORT models. Geosci Remote Sens Lett 14:916–920

Jin Z, Charlock T, Rutledge K, Stamnes K, Wang Y (2006) Analytic solution of radiative transfer in the coupled atmosphere-ocean system with a rough surface. Appl Opt 45:7433–7455

Kotchenova SY, Vermote EF, Matarrese R, Klemm FJ (2006) Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path Radiance Appl Opt 45:6762–6774

Lucht W, Schaaf C, Strahler A (2000) An algorithm for the retrieval of Albedo from space using semi-empirical BRDF models. IEEE Trans Geosci Remote Sens 38:977

Lucht W, Roujean J-L (2000) Considerations in the Parametric modeling of BRDF and Albedo from multiangular satellite sensor observations. Remote Sens Rev 18:343–379

Lyapustin A, Gatebe CK, Kahn R, Brandt R, Redemann J, Russell P, King MD, Pedersen CA, Gerland S, Poudyal R, Marshak A (2010) Analysis of snow bidirectional reflectance from ARCTAS Spring-2008 Campaign. Atmos Chem Phys 10:4359–4375

Mackowski DW, Mishchenko MI (1996) Calculation of the T matrix and the scattering matrix for ensembles of spheres, J Opt Soc Am A 13:2266–2278

Maignan F, Bréon F-M, Fédèle E, Bouvier M (2009) Polarized reflectance of natural surfaces: Spaceborne measurements and analytical modeling. Rem Sens Environ 113:2642–2650

Meng Z, Yang P, Kattawar GW, Bi L, Liou K-N, Lazslo I (2010) Single-scattering properties of triaxial ellipsoidal mineral dust aerosols: A database for application to radiative transfer calculations. J Aer Sci 41:501–512

Mishchenko MI, Travis LD (1998) Capabilities and limitations of a current FORTRAN implementation of the T-matrix method for randomly oriented, rotationally symmetric scatterers, JQSRT 60:309–324

Mishchenko MI, Travis LD (1998) Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight. J Geophys Res 102:16989

Mishchenko MI, Travis LD, Lacis LL (2006) Scattering, absorption and emission of light by small particles. Cambridge University Press, Cambridge, U.K.

Mishchenko MI (2014) Electromagnetic scattering by particles and particle groups: an introduction. Cambridge University Press, ISBN 9780521519922

Morel A, Gentili B (2009) A simple band ratio technique to quantify the colored dissolved and detrital organic material from ocean color remotely sensed data. Remote Sens Environ 113

Nakajima T, Tanaka M (1988) Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation. J Quant Spectrosc Radiat Transf **40**:51–69

Natraj V, Shia RL, Yung YL (2010) On the use of principal component analysis to speed up radiative transfer calculations. J Quant Spectrosc Radiat Transfer 111:810–816, https://doi.org/10.1016/j.jqsrt.2009.11.004

Natraj V, Hovenier JW (2012) Polarized light reflected and transmitted by thick rayleigh scattering atmospheres. Astrophys J 748:28

Rahman H, Pinty B, Verstrate M (1993) Coupled surface-atmospheric reflectance (CSAR) model. 2. Semi-empirical surface model usable with NOAA advanced very high resolution radiometer data. J Geophys Res 98:20791

Rodgers CD (2000) Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing Co. Pte. Ltd., Singapore

Rozanov V, Rozanov A (2007) Relationship between different approaches to derive weighting functions related to atmospheric remote sensing problems. J Quant Spectrosc Radiat Transf 105(2):217–242

Sancer M (1969) Shadow-corrected electromagnetic scattering from a randomly-rough ocean surface. IEEE Trans Antennas Propag **AP-17**:557–585

Sayer A, Thomas G, Grainger R (2010) A sea-surface reflectance model for (A)ATSR, and application to aerosol retrievals. Atmos Meas Tech 3:813–838

Siewert CE (1982) On the phase matrix basic to the scattering of polarized light. Astron Astrophys 109:195–200

Siewert CE (2000a) A concise and accurate solution to Chandrasekhar's basic problem in radiative transfer. J Quant Spectrosc Radiat Transf 64:109–130

Siewert CE (2000b) A discrete-ordinates solution for radiative transfer models that include polarization effects. J Quant Spectrosc Radiat Transf 64:227–254

Spurr R (2002) Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment. J Quant Spectrosc Radiat Transf 75:129–175

Spurr RJD (2004) A New approach to the retrieval of surface properties from earthshine measurements. J Quant Spectrosc Radiat Transf 83:15–46

Spurr RJD (2006) VLIDORT: a linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. J Quant Spectrosc Radiat Transf 102(2):316–342. https://doi.org/10.1016/j.jqsrt.2006.05.005

Spurr R (2008) LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In: Kokhanovsky A (ed) Light scatter reviews, vol 3. Springer, Berlin

Spurr R, Natraj V (2011) A linearized two-stream radiative transfer code for fast approximation of multiple-scatter fields. J Quant Spectrosc Radiat Transfer 112:2630–2637

Spurr R, Christi M (2014) On the generation of atmospheric property Jacobians from the (V)LIDORT linearized radiative transfer models. J Quant Spectrosc Radiat Transf. https://doi.org/10.1016/j.jqsrt.2014.03.011

Spurr R, Christi M (2018) LIDORT-RRS: a linearized discrete ordinate radiative transfer model with first order rotational raman scattering, paper in preparation

Spurr R, Kurosu T, Chance K (2001) A linearized discrete ordinate radiative transfer model for atmospheric remote sensing retrieval. J Quant Spectrosc Radiat Transf 68:689–735

Spurr RJD, de Haan J, van Oss R, Vasilkov A (2008) Discrete ordinate theory in a stratified medium with first order rotational raman scattering; a general quasi-analytic solution. J Quant Spectrosc Radiat Transfer 109:404–425. https://doi.org/10.1016/j.jqsrt.2007.08.011

Spurr R, Wang J, Zeng J, Mishchenko M (2012) Linearized T-Matrix and Mie scattering computations. J Quant Spectrosc Radiat Transfer 113:425–439

Spurr R, Natraj V, Kopparla P, Christi M (2016) Application of principal component analysis (PCA) to performance enhancement of hyperspectral radiative transfer computations, in "Principal Component Analysis: Methods, Applications and Technology", NOVA publishers

Stamnes K, Conklin P (1984) A new multi-layer discrete ordinate approach to radiative transfer in vertically inhomogeneous atmospheres. J Quant Spectrosc Radiat Transfer 31:273

Stamnes K, Tsay S-C, Nakajima T (1988) Computation of eigenvalues and eigenvectors for discrete ordinate and matrix operator method radiative transfer. J Quant Spectrosc Radiat Transf 39:415–419

Ustinov EA (2005) Atmospheric weighting functions and surface partial derivatives for remote sensing of scattering planetary atmospheres in thermal spectral region: General adjoint approach. J Quant Spectrosc Radiat Transf 92:351–371

Van Oss RF, Spurr RJD (2002) Fast and accurate 4 and 6 stream linearized discrete ordinate radiative transfer models for ozone profile retrieval. J Quant Spectrosc Radiat Transf 75:177–220

Vermote EF, Tanré D, Deuzé JL, Herman M, Morcrette JJ (1997) Second simulation of the satellite signal in the solar spectrum, 6S: an overview. IEEE Trans Geosci Remote Sens 35:675–686

Vestrucci M, Siewert CE (1984) A numerical evaluation of an analytical representation of the components in a Fourier decomposition of the phase matrix for the scattering of polarized light. JQSRT 31:177–183

Wanner W, Li X, Strahler A (1995) On the derivation of kernels for kernel-driven models of bidirectional reflectance. J Geophys Res 100:21077

Wiscombe W (1977) The delta-M method: rapid yet accurate radiative flux calculations for strongly asymmetric phase function

Radiative Transfer of Light in Strongly Scattering Media

Boaz Ilan and Arnold D. Kim

1 Introduction

Radiative transfer provides a complete description of absorption, scattering, and radiation of light in a multiple scattering medium (Chandrasekhar 1960; Ishimaru 1999; Van de Hulst 2012). Consequently, radiative transfer is important for several applications such as neutron transport (Bell and Glasstone 1970; Case and Zweifel 1967; Lewis and Miller 1984), astrophysics (Peraiah 2002; Sobolev 2017), geophysics (Tsang et al. 1985; Mobley 1994; Kirk 1994; Thomas and Stamnes 2002; Kokhanovsky 2006b; Marshak and Davis 2005; Mishchenko et al. 2006; Mishchenko 2014), heat transfer (Modest 2013), biomedical optics (Welch et al. 2011; Wang and Wu 2012), and computer graphics (Jensen 2001). Applying radiative transfer theory is challenging for practical problems, because exact solutions of the radiative transfer equation (RTE) are known only for special cases (Case and Zweifel 1967). Even computational methods for radiative transfer require specialized methods (Lewis and Miller 1984; Gao and Zhao 2009) that can be challenging to implement and require substantial computational resources. For these reasons, accurate approximations of solutions of the RTE, and of the RTE itself, are valuable for gaining physical insight.

Boundary layer theory provides a systematic method to solve singularly perturbed boundary value problems (Bender and Orszag 2013; Hinch 1991; Miller 2006). Larsen and Keller (1974) introduced boundary layer theory for general initial-boundary value problems for the RTE. In RTE boundary layer theory, one seeks the asymptotic solution of the RTE in the regime of a strongly-scattering medium. However, it is a rather different approach from that taken in asymptotic radiative transfer

B. Ilan · A. D. Kim (✉)

Applied Mathematics Department, University of California, Merced, 5200 North Lake Road, Merced, CA 95343, USA
e-mail: adkim@ucmerced.edu

B. Ilan
e-mail: bilan@ucmerced.edu

theory (ARTT), which has been reviewed by Kokhanovsky (2006a). In the asymptotic limit of strong multiple scattering, the solution of a boundary value problem for the RTE is written as a sum of an interior solution and boundary layer solutions. The interior solution is governed by the diffusion approximation. The boundary layer solutions correct for errors made by the diffusion approximation near the boundary. Through a systematic perturbation theory, one can determine all of the subproblems required to compute the interior and boundary layer solutions. Moreover, boundary layer theory provides a procedure for computing boundary conditions for the diffusion approximation. The result is a uniformly valid asymptotic approximation to the solution of the RTE that is accurate over the entire domain.

The challenge in applying boundary layer theory to problems of practical interest lies in explicitly computing the boundary layer solutions from which boundary conditions for the diffusion approximation are computed. Habetler and Matkowsky (1975) have applied boundary layer theory to initial-boundary value problems with isotropic scattering in one spatial dimension. For that problem, one can compute the boundary layer solutions explicitly using the singular eigenfunctions described by Case and Zweifel (1967). Malvagi and Pomraning (1991), and Pomraning and Ganapol (1995) have applied boundary layer theory to more general RTE problems by using an approximate variational method to compute the boundary layer solutions. Recently, the authors and their collaborators have used numerical methods in conjunction with boundary layer theory to solve various RTE problems (Kim 2011; Kim and Moscoso 2011; Rohde and Kim 2012, 2014, 2017; Lehtikangas et al. 2012; Şahin-Biryol and Ilan 2014; Dark and Kim 2017). Using numerical methods to compute boundary layer solutions has opened up the use of boundary layer theory to RTE problems of practical interest.

Another approximation, called two-flux theory, is a particularly simple, and intuitive approximation of the RTE in a plane-parallel medium. Its origin traces back to Schuster (1905) in 1905 and Schwarzschild (1906) in 1906. However, it gained popularity due to the works of (Kubelka and Munk 1931; Kubelka 1948). As a result, two-flux theory is often called Kubelka–Munk theory. This theory has found extensive use for various applications (e.g., see the review by Philips-Invernizzi et al. (2001), largely due to its simplicity.

Two-flux theory models light fluxes traveling in forward and backward directions. Absorption and scattering in the medium are quantified using absorption and scattering coefficients, typically denoted by K, and S, respectively. One major challenge in two-flux theory lies in connecting these coefficients to fundamental quantities in radiative transfer. Many studies (Brinkworth 1972; Gate 1974; Nobbs 1985; Star et al. 1988; Vargas and Niklasson 1997; Molenaar et al. 1999; Yang and Kruse 2004; Yang et al. 2004; Yang and Miklavcic 2005; Edström 2007; Thennadil 2008; Neuman and Edström 2010; Myrick et al. 2011; Kokhanovsky 2007) have addressed the theoretical basis for two-flux theory and its applicability to interpret measured data. Sandoval and Kim (2014) derive two-flux theory using a systematic perturbation analysis of the double-spherical harmonics method of order one in the limit of strong multiple scattering. This perturbation method is similar to what is done in boundary layer theory.

In this paper, we review boundary layer theory and two-flux theory and draw connections between them. In particular, we derive two-flux theory using the interior solution from boundary layer theory governed by the diffusion approximation. This derivation is more elementary than the one done by Sandoval and Kim (2014). Consequently, it provides the framework needed to understand that two-flux theory is simply a reformulation of the diffusion approximation. We use the double-spherical harmonics method (Case and Zweifel 1967; Lewis and Miller 1984; Aronson 1986) to compute boundary layer solutions which, in turn, provide an efficient and effective method for computing boundary conditions for the diffusion approximation. We use the boundary conditions for the diffusion approximation to derive boundary conditions for two-flux theory. We call this model the asymptotic two-flux theory, because it is asymptotically accurate for strongly-scattering media. In addition, by considering the double-spherical harmonics method of order one, we draw its equivalence to four-flux theory and its connection to boundary layer theory. Through these connections, we develop valuable insights into radiative transfer problems in strongly-scattering media.

By comparing results by boundary layer theory and two-flux theory with direct numerical solutions of the full RTE, we show their respective scope of applicability. Boundary layer theory gives a uniformly-valid asymptotic approximation meaning that it is accurate throughout the entire domain with an a-priori error bound. In contrast, two-flux theory is not a uniformly valid asymptotic approximation– its accuracy depends on the boundary conditions prescribed. For standard boundary conditions based on the in-going flux at the boundary, two-flux theory is accurate near the boundaries, but inaccurate in the interior of the medium. When boundary conditions for two-flux theory are prescribed in a manner consistent with boundary layer theory, we find that it is accurate in the interior of the medium, but not near the boundaries. Both boundary layer and two-flux theories are valid for strongly scattering media where scattering is not too sharply forward peaked.

The remainder of this paper is as follows. In Sect. 2 we discuss the RTE in a plane-parallel slab, and describe all of the physical quantities in that problem. In Sect. 3, we give explicit results for the spectrum of the scattering operator in the RTE. These results are used throughout the analysis that follows. We describe boundary layer theory for the RTE in Sect. 4 including all of the details in deriving the interior and boundary layer solutions. In Sect. 5 we describe the double-spherical harmonics method used to compute the boundary layer solutions which, in turn, provides a method to compute boundary conditions for the diffusion approximation governing the interior solution. We summarize the results from boundary layer theory using the double-spherical harmonics method in a procedure to compute the asymptotic approximation in Sect. 6. In Sect. 7 we give a new derivation of two-flux theory based on the interior solution from boundary layer theory. In doing so, we are able to identify its limitations and prescribe boundary conditions consistent with boundary layer theory leading to the asymptotic two-flux theory. We show computational results comparing all of these approximations in Sect. 8. Additionally, Sect. 8 discusses the

relationship between four-flux theory and the double-spherical harmonics method of order one. We discuss the extension of two-flux theory for the RTE in three dimensions in Sect. 9. Section 10 gives our conclusions.

2 Radiative Transfer in a Plane-Parallel Slab

Consider a uniformly absorbing and scattering medium contained in a plane-parallel slab. When a continuous source of light penetrates into this medium, be it from incident radiation and/or emanating from the surface, the light's radiance (or specific intensity), $I(\boldsymbol{x}, \boldsymbol{\Omega})$, can be described by the radiative transfer equation (RTE) (Chandrasekhar 1960),

$$\boldsymbol{\Omega} \cdot \nabla I + \kappa_a I + \kappa_s L I = 0. \tag{1}$$

Here, the spatial coordinates parallel to the slab surface are $(x, y) \in \mathbb{R}^2$, and the coordinate normal to the slab surface is $z \in (z_1, z_2)$. The direction cosines Ω_1, Ω_2, and Ω_3 are the components of $\boldsymbol{\Omega}$, which is a vector on the unit sphere S^2. The absorption and scattering coefficients are denoted by κ_a and κ_s, respectively. The scattering operator L is defined by

$$L I = I - \int_{S^2} p(\boldsymbol{\Omega}, \boldsymbol{\Omega}') I(\boldsymbol{\Omega}') \, d\boldsymbol{\Omega}, \tag{2}$$

where $p(\cdot)$ is called the scattering phase function. In a bounded domain, boundary conditions prescribe the radiance for all directions pointing into the domain (Case and Zweifel 1967; Ishimaru 1999). In general, boundary conditions take the form

$$I|_{z=z_1} = b_1 \quad \text{on} \quad \{\boldsymbol{x} \in \mathbb{R}^2, \boldsymbol{\Omega} \cdot \hat{\boldsymbol{z}} > 0\}, \tag{3a}$$

$$I|_{z=z_2} = b_2 \quad \text{on} \quad \{\boldsymbol{x} \in \mathbb{R}^2, \boldsymbol{\Omega} \cdot \hat{\boldsymbol{z}} < 0\}, \tag{3b}$$

where $b_1(\boldsymbol{x}, \boldsymbol{\Omega})$ and $b_2(\boldsymbol{x}, \boldsymbol{\Omega})$ prescribe the radiance entering the medium at $z = z_1$ and $z = z_2$, respectively, and $\hat{\boldsymbol{z}}$ denotes the unit vector pointing along the z-axis in the positive sense. We can consider boundary conditions that take into account reflections due to a refractive index mismatch. All of the results that follow can take these reflecting boundary conditions into account. However, they complicate the analysis without providing any additional insight, so we do not consider them here.

We assume axi-symmetry so that $b_1 = b(\mu)$ where $\mu = \cos\theta$ and θ is the inclination angle made with respect to the z-axis. Moreover, we assume the radiance enters the medium on $z = z_1$ uniformly in (x, y). That light is the only source in the problem, so that $b_2 = 0$ on the surface $z = z_2$. Furthermore, we assume that the scattering is rotationally invariant and depends only on the angle made between the

incident and scattering directions, so that $p = p(\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}')$. In this case, the solution of this boundary-value problem is uniform in (x, y) and azimuthally symmetric, i.e., $I = I(z, \mu)$. For this case, I is governed by the *one-dimensional* RTE

$$\mu \frac{\mathrm{d}I}{\mathrm{d}z} + \kappa_a I + \kappa_s \tilde{L} I = 0 , \tag{4}$$

where the azimuthally-integrated scattering operator is (henceforth, the tilde on $\tilde{L}$ is dropped)

$$L I = I - \frac{1}{2} \int_{-1}^{1} h(\mu, \mu') I(z, \mu') \, \mathrm{d}\mu' . \tag{5}$$

The redistribution function h in (4) is defined in terms of the scattering phase function as

$$h(\mu, \mu') = \frac{1}{2\pi} \int_{0}^{2\pi} p\left(\mu\mu' + \sqrt{1 - \mu^2}\sqrt{1 - \mu'^2} \cos(\varphi - \varphi') \right) \mathrm{d}(\varphi - \varphi') \tag{6}$$

and normalized according to

$$\frac{1}{2} \int_{-1}^{1} h(\mu, \mu') \mathrm{d}\mu' = 1 . \tag{7}$$

Equation (4) is to be solved in $z_1 < z < z_2$ subject to the boundary conditions

$$I|_{z=z_1} = b(\mu) \quad \text{on } 0 < \mu \leq 1, \tag{8a}$$
$$I|_{z=z_2} = 0 \quad \text{on } -1 \leq \mu < 0. \tag{8b}$$

It is convenient to define the non-dimensional distance variable, called the optical depth,

$$\tau = (\kappa_a + \kappa_s) z . \tag{9}$$

Making the transformation of variable from z to τ and introducing the albedo,

$$\varpi_0 = \frac{\kappa_s}{\kappa_s + \kappa_a} , \tag{10}$$

we find that (4) becomes

$$\mu \frac{\mathrm{d}I}{\mathrm{d}\tau} + I = \frac{\varpi_0}{2} \int_{-1}^{1} h(\mu, \mu') I(z, \mu') . \tag{11}$$

The albedo defined in (10) satisfies $0 < \varpi_0 \leq 1$, with $\varpi_0 = 1$ corresponding to *conservative scattering* in which $\kappa_a = 0$.

3 Legendre Polynomial Expansion and the Spectrum of the Scattering Operator

The Legendre polynomials are a sequence of orthogonal polynomials, $P_1(\mu)$, $P_2(\mu), \ldots$. They arise as a special case of spherical harmonics (Courant and Hilbert 2008) for azimuthally invariant functions. These polynomials satisfy the orthogonality relation

$$\frac{1}{2} \int_{-1}^{1} P_m(\mu) P_n(\mu) \mathrm{d}\mu = \frac{1}{2n+1} \delta_{mn}, \tag{12}$$

with δ_{mn} denoting the Kronecker delta. The sequence $\{P_n(\nu)\}$, $n = 0, 1, 2, \ldots$ is complete in the L_2 sense. In our analysis we shall frequently refer to the first two Legendre polynomials, which are $P_0(\mu) = 1$ and $P_1(\mu) = \mu$. We shall also make use of the recursion relation

$$\mu P_n(\mu) = \frac{1}{2n+1} \left[(n+1) P_{n+1}(\mu) + n P_{n-1}(\mu) \right]. \tag{13}$$

Because the redistribution function (6) is inherently rotationally invariant, it can be expanded in Legendre polynomials as

$$h(\mu, \mu') = \sum_{n=0}^{\infty} (2n+1) g_n P_n(\mu) P_n(\mu'), \tag{14}$$

where the coefficients are

$$g_n = \frac{2n+1}{2} \int_{-1}^{1} \left[\frac{1}{2} \int_{-1}^{1} h(\mu, \mu') P_n(\mu') \mathrm{d}\mu' \right] P_n(\mu) \mathrm{d}\mu. \tag{15}$$

Because the scattering phase function is normalized according to (7), it follows that the first coefficient of (15) is $g_0 = 1$. The second coefficient g_1 is called the anisotropy factor or mean cosine, which we denote by $g_1 = g$. It is given by

$$g = \frac{3}{4} \int_{-1}^{1} \left[\int_{-1}^{1} h(\mu, \mu') \mu \mu' \mathrm{d}\mu' \right] \mathrm{d}\mu. \tag{16}$$

The spectrum of the operator L, given in (5), plays a key role in our analysis. In light of (14), we find that

$$L P_n(\mu) = (1 - g_n) P_n(\mu). \tag{17}$$

It follows that $P_n(\mu)$ is an eigenfunction of L with eigenvalue $1 - g_n$. For the standard redistribution functions considered in the literature, all the eigenvalues are simple and decreasing, i.e., $g_n > g_{n+1}$ for all n. Since $g_0 = 1$, the lowest eigenvalue of L

is zero. All the other eigenvalues are positive because $g_n < 1$ for $n > 0$ (Sobolev 2017; Kim and Keller 2003). To summarize, L is a non-negative operator and the isotropic function, $P_0(\mu) = 1$, spans its nullspace. The spectrum asymptotes to one since $g_n \to 0$ as $n \to \infty$. It also follows from (16) that the first non-zero eigenvalue of L is $1 - g$, i.e.,

$$L\mu = (1 - g)\mu . \tag{18}$$

4 Boundary Layer Theory and the Diffusion Approximation

Boundary layer theory for radiative transfer is an asymptotic theory for highly-scattering and weakly-absorbing media. There are three characteristic length scales in (4): the characteristic absorption length, $\ell_a = \kappa_a^{-1}$, the scattering mean-free path, $\ell_s = \kappa_s^{-1}$, and the thickness of the slab, $\ell_z = z_2 - z_1$. The highly-scattering and weakly-absorbing regime corresponds to $\ell_s \ll \ell_z \ll \ell_a$. To make this precise, we introduce the small, dimensionless parameter $0 < \varepsilon \ll 1$ and set

$$\kappa_s \ell_z = \varepsilon^{-1} , \tag{19a}$$

$$\kappa_a \ell_z = \varepsilon\alpha, \tag{19b}$$

where α is a rescaled absorption coefficient. We make the change of variables

$$z \to \tilde{z} = \frac{z - z_1}{\ell_z} . \tag{20}$$

Henceforth, the tilde sign will be dropped. Substituting (20), (19a), and (19b) into (4), we obtain

$$\varepsilon\mu\frac{dI}{dz} + \varepsilon^2\alpha I + LI = 0. \tag{21}$$

Equation (21) is to be solved in $0 < z < 1$ subject to boundary conditions

$$I|_{z=0} = b(\mu) \quad \text{on} \quad 0 < \mu \le 1, \tag{22a}$$

$$I|_{z=1} = 0 \quad \text{on} \; -1 \le \mu < 0. \tag{22b}$$

We remark that in terms of (19a) and (19b), the single scattering albedo (10) is

$$\varpi_0 = \frac{\kappa_s}{\kappa_s + \kappa_a} = \frac{1}{1 + \varepsilon^2\alpha} = 1 - \varepsilon^2\alpha + \mathcal{O}(\varepsilon^4). \tag{23}$$

Thus, the albedo for this scaling is near unity, i.e., $1 - \varpi_0 = \mathcal{O}(\varepsilon^2)$. It is well-known that the diffusion approximation is valid in this regime. The diffusion approximation asserts that the radiance becomes nearly isotropic deep inside the medium due to

strong multiple scattering. In what follows, we derive the diffusion approximation using perturbation methods.

We seek the solution of (21) subject to boundary conditions (22) in the asymptotic limit as $\varepsilon \to 0^+$. If we were to naïvely take the limit of (21) as $\varepsilon \to 0^+$, the derivative term in (21) vanishes and the resulting equation could not satisfy boundary conditions (8). In fact, the solution of this boundary value problem rapidly changes near the boundaries $z = 0$ and $z = 1$. For this reason, (21) is called a singularly perturbed equation. To address this, we seek its solution as the sum

$$I = \Phi + \Psi^{(1)} + \Psi^{(2)}, \tag{24}$$

with Φ denoting the interior solution and $\Psi^{(1)}$ (resp. $\Psi^{(2)}$) denoting the boundary layer solution near $z = 0$ (resp. $z = 1$). In what follows, we seek to find Φ, $\Psi^{(1)}$, and $\Psi^{(2)}$ using perturbation methods.

4.1 Interior Solution

We seek the asymptotic expansion for the interior solution in the form

$$\Phi \sim \sum_{n=0}^{\infty} \varepsilon^n \phi_n, \quad \varepsilon \to 0^+. \tag{25}$$

Substituting (25) into (21) and collecting like-powers of ε, we find that

$$\mathcal{O}(1): \quad L\phi_0 = 0, \tag{26a}$$

$$\mathcal{O}(\varepsilon): \quad L\phi_1 = -\mu \frac{d\phi_0}{dz}, \tag{26b}$$

$$\mathcal{O}(\varepsilon^n): \quad L\phi_n = -\mu \frac{d\phi_{n-1}}{dz} - \alpha\phi_{n-2}, \quad n \geq 2. \tag{26c}$$

We recall from Sect. 3 that $P_0(\mu)$ spans the nullspace of L. Therefore, the general solution of (26a) is

$$\phi_0 = P_0(\mu)\rho_0(z), \tag{27}$$

where $\rho_0(z)$ is an undetermined function of z. Substituting (27) into (26b) and recalling that $P_1(\mu) = \mu$ gives

$$L\phi_1 = -\mu\rho_0'(z) = -P_1(\mu)\rho_0'(z). \tag{28}$$

Equation (28) is a linear, non-homogeneous Fredholm integral equation of the second kind. Its solution can be decomposed as the sum of the homogeneous solution ϕ_1^H and the particular solution ϕ_1^P. Following the analysis of (26a), $\phi_1^H = \rho_1(z)P_0(\mu)$

where $\rho_1(z)$ is an undetermined function of z. The condition for the solvability of the nonhomogeneous problem, $L\phi = f$, with f denoting a general nonhomogeneous term, is given by

$$\int_0^1 P_0(\mu) f(\mu)\mathrm{d}\mu = \int_0^1 f(\mu)\mathrm{d}\mu = 0. \tag{29}$$

In other words, the right-hand side of (28) (i.e., the non-homogeneous term) must be orthogonal to the nullspace of L. Because $P_1(\mu)$ is orthogonal to $P_0(\mu)$ [see (12)], in this case, solvability condition (29) is automatically satisfied. Recall that $LP_1(\mu) = (1-g)P_1(\mu)$. Substituting $\phi_1^P = CP_1(\mu)\rho_0'(z)$ into (28), we find that $C = -1/(1-g)$. It follows that

$$\phi_1(z,\mu) = \rho_1(z)P_0(\mu) - \frac{1}{1-g}\rho_0'(z)P_1(\mu). \tag{30}$$

Substituting (27) and (30) into (26c) with $n = 2$, we obtain

$$L\phi_2 = -\mu\rho_1'(z)P_0(\mu) + \frac{1}{1-g}\rho_0''(z)\mu P_1(\mu) - \alpha\rho_0(z)P_0(\mu) . \tag{31}$$

Applying the recursion relation (13) with $n = 1$ yields

$$L\phi_2 = -\rho_1'(z)P_1(\mu) + \frac{1}{3(1-g)}\rho_0''(z)[2P_2(\mu) + P_0(\mu)] - \alpha\rho_0(z)P_0(\mu) . \tag{32}$$

Applying solvability condition (29) to the right-hand side of (32) leads to

$$\frac{1}{3(1-g)}\rho_0''(z) - \alpha\rho_0(z) = 0 . \tag{33}$$

Equation (33) is called the diffusion equation. It gives the leading order asymptotic behavior of the radiance in the interior of the medium.

It is instructive and useful to obtain higher-order corrections to the diffusion approximation. To this end, we substitute (33) into (32) and obtain

$$L\phi_2 = -\rho_1'(z)P_1(\mu) + \frac{2}{3(1-g)}\rho_0''(z)P_2(\mu) . \tag{34}$$

From spectral equation (17), we have $LP_1 = (1-g)P_1$ and $LP_2 = (1-g_2)P_2$. Substituting these into (34) we find that

$$\phi_2(z,\mu) = \rho_2(z)P_0(\mu) - \frac{1}{1-g}\rho_1'(z)P_1(\mu) + \frac{2}{3(1-g)(1-g_2)}\rho_0''(z)P_2(\mu) , \tag{35}$$

with $\rho_2(z)$ an undetermined function of z. Substituting (30) and (35) into (26c) with $n = 3$, we find that

$$L\phi_3 = -\rho_2'(z)\mu + \frac{1}{1-g}\rho_1''(z)\mu P_1(\mu) - \frac{2}{3(1-g)(1-g_2)}\rho_0'''(z)\mu P_2(\mu)$$
$$- \alpha\rho_1(z)P_0(\mu) + \frac{\alpha}{1-g}\rho_0'(z)P_1(\mu) .$$

$$(36)$$

Applying the recursion relation (13) with $n = 1$ and $n = 2$ leads to

$$L\phi_3 = -\rho_2'(z)P_1(\mu) + \frac{1}{3(1-g)}\rho_1''(z)\left[2P_2(\mu) + P_0(\mu)\right]$$
$$- \frac{2}{15(1-g)(1-g_2)}\rho_0'''(z)\left[3P_3(\mu) + 2P_1(\mu)\right]$$
$$- \alpha\rho_1(z)P_0(\mu) + \frac{\alpha}{1-g}\rho_0'(z)P_1(\mu).$$

$$(37)$$

Applying solvability condition (29) to the right-hand side of (37) yields

$$\frac{1}{3(1-g)}\rho_1''(z) - \alpha\rho_1(z) = 0 .$$

$$(38)$$

Remarkably, (38) is the same as (33). We can continue on to determine ϕ_3. The result will be a linear combination of $P_3(\mu)$, $P_2(\mu)$, and $P_1(\mu)$ with coefficients involving derivatives of ρ_1 and ρ_0 and the homogeneous solution, $P_0(\mu)\rho_3(z)$, with $\rho_3(z)$ an undetermined function of z. By substituting that result into the equation for ϕ_4 and ensuring its solvability, we will find that ρ_2 satisfies the same diffusion equation as ρ_0 and ρ_1.

Thus far, we have obtained an $\mathcal{O}(\varepsilon^2)$ approximation of the interior solution as

$$\Phi(z, \mu) = \rho_0(z)P_0(\mu) + \varepsilon\left[\rho_1(z)P_0(\mu) - \frac{\rho_0'(z)P_1(\mu)}{1-g}\right]$$
$$+ \varepsilon^2\left[\rho_2(z)P_0(\mu) - \frac{\rho_1'(z)P_1(\mu)}{1-g} + \frac{2\rho_0''P_2(\mu)}{3(1-g)(1-g_2)}\right] + O(\varepsilon^3), \quad (39)$$

with ρ_0, ρ_1, and ρ_2 satisfying the diffusion equations

$$\frac{1}{3(1-g)}\rho_n''(z) - \alpha\rho_n(z) = 0, \quad n = 0, 1, 2 .$$

$$(40)$$

These diffusion equations are to be solved in $0 < z < 1$. However, this perturbation analysis does not provide a method to determine boundary conditions for (40). To address this, we analyze the boundary layer solutions.

4.2 Boundary Layer Solutions

The boundary layer solution $\Psi^{(1)}$ (resp. $\Psi^{(2)}$) is a special solution of (21) that decays rapidly in a region of thickness $\mathcal{O}(\varepsilon)$ about $z = 0$ (resp. $z = 1$), called the boundary layer, and is exponentially small outside of this region.

To determine $\Psi^{(1)}$, we introduce the stretched variable $z = \varepsilon Z$ and set $\Psi^{(1)}(\mu, Z) = I(\mu, \varepsilon Z)$. Rewriting (21) in terms of Z and $\Psi^{(1)}$, we obtain

$$\mu \frac{d\Psi^{(1)}}{dZ} + L\Psi^{(1)} = -\varepsilon^2 \alpha \Psi^{(1)}. \tag{41}$$

Substituting (24) into the boundary condition (8a), neglecting terms involving $\Psi^{(2)}$, and solving for $\Psi^{(1)}$ leads to

$$\Psi^{(1)}|_{Z=0} = b - \Phi|_{z=0} \quad \text{on} \quad 0 < \mu \le 1. \tag{42}$$

We are able to neglect terms involving $\Psi^{(2)}$ in (42) because it is constructed to decay rapidly in a region of thickness $\mathcal{O}(\varepsilon)$ about $z = 1$. Hence, it will be exponentially small near $Z = 0$. Using (39) to $\mathcal{O}(\varepsilon)$ yields

$$\Psi^{(1)}|_{Z=0} = b - \rho_0(0) - \varepsilon \left[\rho_1(0) - \frac{1}{1-g} \mu \rho_0'(0) \right] \quad \text{on} \quad 0 < \mu \le 1. \tag{43}$$

However, to ensure asymptotic matching with the interior solution, we also require that

$$\Psi^{(1)} \to 0 \quad \text{as} \quad Z \to \infty, \tag{44}$$

which is called the asymptotic matching condition or the radiation condition. Thus, (41) is to be solved in the half space, $Z > 0$, subject to boundary condition (42) and radiation condition (44). We seek an asymptotic solution of the form

$$\Psi^{(1)} \sim \sum_{n=0}^{\infty} \varepsilon^n \psi_n^{(1)}, \quad \varepsilon \to 0^+. \tag{45}$$

Substituting (45) into (41), (43), and (44), and collecting like-powers of ε, we obtain at $\mathcal{O}(1)$

$$\mu \frac{d\psi_0^{(1)}}{dZ} + L\psi_0^{(1)} = 0 \quad \text{in} \quad Z > 0 \tag{46a}$$

$$\psi_0^{(1)}|_{Z=0} = b - \rho_0(0) \quad \text{on} \quad 0 < \mu \le 1, \tag{46b}$$

$$\psi_0^{(1)} \to 0 \quad \text{as} \quad Z \to \infty, \tag{46c}$$

and at $\mathcal{O}(\varepsilon)$

$$\mu \frac{d\psi_1^{(1)}}{dZ} + L\psi_1^{(1)} = 0 \quad \text{in} \quad Z > 0 \tag{47a}$$

$$\psi_1^{(1)}|_{Z=0} = -\rho_1(0) + \frac{1}{1-g}\mu\rho_0'(0) \quad \text{on} \quad 0 < \mu \le 1, \tag{47b}$$

$$\psi_1^{(1)} \to 0 \quad \text{as} \quad Z \to \infty, \tag{47c}$$

The problems (46) and (47) determine the boundary layer solution in the boundary layer near $z = 0$ up to $\mathcal{O}(\varepsilon^2)$.

Similarly, we can derive corresponding problem for the boundary layer solution $\Psi^{(2)}$ in the boundary layer near $z = 1$. For that case, we introduce the stretched variable $z = 1 - \varepsilon\tilde{Z}$ and set $\Psi^{(2)}(\mu, \tilde{Z}) = I(\mu, 1 - \varepsilon\tilde{Z})$ which satisfies

$$-\mu \frac{d\Psi^{(2)}}{d\tilde{Z}} + L\Psi^{(2)} = -\varepsilon^2\alpha\Psi^{(2)}. \tag{48}$$

Substituting (24) into the boundary condition (8b), neglecting terms involving $\Psi^{(1)}$, and solving for $\Psi^{(2)}$ leads to

$$\Psi^{(2)}|_{\tilde{Z}=0} = -\Phi|_{z=1} \quad \text{on} - 1 \le \mu < 0. \tag{49}$$

Using (39) to $\mathcal{O}(\varepsilon)$ yields

$$\Psi^{(2)}|_{\tilde{Z}=0} = -\rho_0(1) - \varepsilon\left[\rho_1(1) - \frac{1}{1-g}\mu\rho_0'(1)\right] \quad \text{on} - 1 \le \mu < 0. \tag{50}$$

We require for asymptotic matching that

$$\Psi^{(2)} \to 0 \quad \text{as} \quad \tilde{Z} \to \infty . \tag{51}$$

We seek an asymptotic solution of the form

$$\Psi^{(2)} \sim \sum_{n=0}^{\infty} \varepsilon^n \psi_n^{(2)}, \quad \varepsilon \to 0^+, \tag{52}$$

and find that at $\mathcal{O}(1)$

$$-\mu \frac{d\psi_0^{(2)}}{d\tilde{Z}} + L\psi_0^{(2)} = 0 \quad \text{in} \quad \tilde{Z} > 0, \tag{53a}$$

$$\psi_0^{(2)}|_{\tilde{Z}=0} = -\rho_0(1) \quad \text{on} - 1 \le \mu < 0, \tag{53b}$$

$$\psi_0^{(2)} \to 0 \quad \text{as} \quad \tilde{Z} \to \infty, \tag{53c}$$

and at $\mathscr{O}(\varepsilon)$

$$-\mu \frac{d\psi_1^{(2)}}{d\tilde{Z}} + L\psi_1^{(2)} = 0 \quad \text{in} \quad \tilde{Z} > 0, \tag{54a}$$

$$\psi_1^{(2)}|_{\tilde{Z}=0} = -\rho_1(1) + \frac{1}{1-g}\mu\rho_0'(1) \quad \text{on} \quad -1 \le \mu < 0, \tag{54b}$$

$$\psi_1^{(2)} \to 0 \quad \text{as} \quad \tilde{Z} \to \infty. \tag{54c}$$

4.3   Boundary Conditions for the Diffusion Approximation

The boundary layer solutions satisfy conservative (no absorption), one-dimensional RTEs. These problems cannot be solved analytically, in general. We will show how to compute them numerically in the following section using the double-spherical harmonics method. However, let us suppose here that we have the surface Green's function $G_s(Z, \mu; \mu')$ satisfying

$$\mu \frac{dG_s}{dZ} + LG_s = 0 \quad \text{in} \quad Z > 0 \tag{55a}$$

$$G_s|_{Z=0} = \delta(\mu - \mu') \quad \text{on} \quad 0 < \mu, \mu' \le 1, \tag{55b}$$

Using this surface Green's function, the solution of

$$\mu \frac{d\psi}{dZ} + L\psi = 0 \quad \text{in} \quad Z > 0, \tag{56a}$$

$$\psi|_{Z=0} = b(\mu) \quad \text{on} \quad 0 < \mu \le 1, \tag{56b}$$

is given as

$$\psi(Z, \mu) = (\mathscr{G}_s b)(Z, \mu) = \int_0^1 G_s(Z, \mu; \mu')b(\mu')d\mu'. \tag{57}$$

While the solution given in (57) satisfies (56a) and boundary condition (56b), it does not satisfy an asymptotic matching condition of the form $\psi \to 0$ as $Z \to \infty$. We observe that a $\psi = $ constant is a solution of (56a). This constant mode is the only one that does not vanish as $Z \to \infty$. Therefore, we must remove the constant mode from the solution to ensure asymptotic matching. Let the operator $\mathscr{G}_s^0$ be the part of $\mathscr{G}_s$ that maps the boundary data b to the constant mode. To ensure asymptotic matching, we must have the boundary data satisfy

$$(\mathscr{G}_s^0 b) = 0. \tag{58}$$

Let

$$\mathscr{G}_s^0[1] = \beta_0, \tag{59}$$

$$\mathscr{G}_s^0[\mu] = \beta_1, \tag{60}$$

$$\mathscr{G}_s^0[b(\mu)] = \bar{b}. \tag{61}$$

It follows from applying $\mathscr{G}_s^0$ to the right-hand side of (46b) and setting that result to zero that

$$\mathscr{G}_s^0[b - \rho_0(0)] = \bar{b} - \beta_0 \rho_0(0) = 0. \tag{62}$$

This result sets the boundary condition for ρ_0 on $z = 0$. Doing the same procedure on (47b), we obtain

$$\mathscr{G}_s^0 \left[-\rho_1(0) + \mu \frac{1}{1-g} \rho_0'(0) \right] = -\beta_0 \rho_1(0) + \frac{\beta_1}{1-g} \rho_0'(0) = 0, \tag{63}$$

which sets a boundary condition for $\rho_1(0)$.

Let $\tilde{G}_s(Z, \mu; \mu')$ satisfy

$$-\mu \frac{\mathrm{d}\tilde{G}_s}{\mathrm{d}\tilde{Z}} + L\tilde{G}_s = 0 \quad \text{in} \quad \tilde{Z} > 0 \tag{64a}$$

$$\tilde{G}_s|_{\tilde{Z}=0} = \delta(\mu - \mu') \quad \text{on} \quad -1 \le \mu, \mu' < 0. \tag{64b}$$

Then the solution of

$$-\mu \frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}\tilde{Z}} + L\psi = 0 \quad \text{in} \quad \tilde{Z} > 0, \tag{65a}$$

$$\tilde{\psi}|_{\tilde{Z}=0} = \tilde{b}(\mu) \quad \text{on} \quad -1 \le \mu < 0, \tag{65b}$$

is given as

$$\tilde{\psi}(Z, \mu) = (\tilde{\mathscr{G}}_s \tilde{b})(Z, \mu) = \int_{-1}^{0} \tilde{G}_s(Z, \mu; \mu') \tilde{b}(\mu') \mathrm{d}\mu'. \tag{66}$$

In fact, $\tilde{G}_s(\tilde{Z}, \mu; \mu') = G_s(\tilde{Z}, -\mu; -\mu')$. Let $\tilde{\mathscr{G}}_s^0$ be defined for $\tilde{G}_s$ analogously to how $\mathscr{G}_s^0$ is defined for G_s. It follows from applying $\tilde{\mathscr{G}}_s^0$ to the right-hand side of (53b) and setting that result to zero, we obtain

$$\tilde{\mathscr{G}}_s^0[-\rho_0(1)] = -\beta_0 \rho_0(1) = 0. \tag{67}$$

This result sets a boundary condition for ρ_0 on $z = 1$. Doing the same procedure on (54b), we find that

$$\tilde{\mathscr{G}}_s^0 \left[-\rho_1(1) + \mu \frac{1}{1-g} \rho_0'(1) \right] = -\beta_0 \rho_1(1) - \frac{\beta_1}{1-g} \rho_0'(0) = 0, \qquad (68)$$

which sets a boundary condition for ρ_1 on $z = 1$.

Through this analysis of the boundary layer solutions, we are able to determine boundary conditions for the diffusion approximation. To summarize these results, we have

$$\rho_0'' - 3\alpha(1-g)\rho_0 = 0, \quad \text{in} \ \ 0 < z < 1, \qquad (69a)$$

$$\beta_0 \rho_0(0) = \bar{b}, \qquad \beta_0 \rho_0(1) = 0, \qquad (69b)$$

and

$$\rho_1'' - 3\alpha(1-g)\rho_1 = 0, \quad \text{in} \ \ 0 < z < 1, \qquad (70a)$$

$$\beta_0 \rho_1(0) = \beta_1 \frac{\rho_0'(0)}{1-g}, \qquad \beta_0 \rho_1(1) = -\beta_1 \frac{\rho_0'(1)}{1-g}, \qquad (70b)$$

Since ρ_0 and ρ_1 satisfy the same diffusion equation, we introduce $\rho = \rho_0 + \varepsilon\rho_1$ which satisfies

$$\rho'' - 3\alpha(1-g)\rho = 0, \quad \text{in} \ \ 0 < z < 1, \qquad (71a)$$

$$\beta_0 \rho(0) - \varepsilon\beta_1 \frac{\rho'(0)}{1-g} = \bar{b}, \qquad \beta_0 \rho(1) + \varepsilon\beta_1 \frac{\rho'(1)}{1-g} = 0. \qquad (71b)$$

This form satisfies the diffusion equation with Robin boundary conditions that are typically used for the diffusion approximation. Here, the coefficients, β_0 and β_1 and boundary data $\bar{b}$ are computed from the boundary layer problems defined above to ensure asymptotic matching.

5 Double-Spherical Harmonics Method

The double-spherical harmonics method is an efficient numerical method for computing solutions of the RTE (Case and Zweifel 1967; Lewis and Miller 1984; Aronson 1986). We discuss this method for solving boundary layer problems: (46), (47), (53), and (54). By doing so, we obtain a method to compute the coefficients needed for the boundary conditions of the diffusion approximation. In particular, we study the conservative, one-dimensional RTE

$$\mu \frac{d\psi}{dZ} + L\psi = 0 \ \ \text{in} \ \ Z > 0, \qquad (72)$$

subject to boundary condition

$$\psi|_{Z=0} = b \quad \text{on} \quad 0 < \mu \leq 1. \tag{73}$$

We will consider conditions required to satisfy the asymptotic matching condition:

$$\psi \to 0, \quad \text{as} \quad Z \to \infty. \tag{74}$$

The solution of (72) is discontinuous for $\mu = 0$ on $Z = 0$, i.e., on the boundary along the direction that is tangent to the boundary. This discontinuity decays exponentially away from the boundary due to scattering and absorption, so it does not affect the well-posedness of this problem. However, its presence introduces a computational difficulty. We address this issue by introducing the half-range specific intensities,

$$\psi_{\pm}(Z, \mu) = \psi(Z, \pm\mu), \quad 0 < \mu \leq 1. \tag{75}$$

Substituting (75) into (72) and (5) leads to the coupled system

$$\mu \frac{d\psi_+}{dZ} = -\psi_+ + H_1\psi_+ + H_2\psi_-, \tag{76a}$$

$$-\mu \frac{d\psi_-}{dZ} = -\psi_- + H_1\psi_- + H_2\psi_+ \tag{76b}$$

with the boundary condition

$$\psi_+|_{Z=0} = b \quad \text{on} \quad 0 < \mu \leq 1. \tag{77}$$

In (76) the half-range integral operators: H_1 and H_2, are defined by

$$H_1\psi_{\pm} = \frac{1}{2} \int_0^1 h(\mu, \mu')\psi_{\pm}(Z, \mu')d\mu', \tag{78a}$$

$$H_2\psi_{\pm} = \frac{1}{2} \int_0^1 h(\mu, -\mu')\psi_{\pm}(Z, \mu')d\mu'. \tag{78b}$$

Since scattering is assumed to be rotationally invariant (see (6)), it follows that $h(-\mu, -\mu') = h(\mu, \mu')$ and $h(\mu, -\mu') = h(-\mu, \mu')$. These symmetries have been used in deriving (76).

Let

$$\tilde{P}_n(\mu) = \sqrt{2n + 1}\, P_n(2\mu - 1) \tag{79}$$

denote the normalized Legendre polynomial mapped to the half range, $0 < \mu \leq 1$. For the double-spherical harmonics method, we expand $\psi^{\pm}$ as

$$\psi_{\pm}(z, \mu) = \sum_{n=0}^{\infty} u_n^{\pm}(z) \tilde{P}_n(\mu) . \tag{80}$$

The objective is to determine the coefficients $u_n^{\pm}(z)$. Substituting (80) into (76), multiplying this by $\tilde{P}_m(\mu)$, integrating with respect to μ over $0 < \mu \leq 1$, and using the half-space orthogonality relation,

$$\int_0^1 \tilde{P}_m(\mu) \tilde{P}_n(\mu) \mathrm{d}\mu = \delta_{mn} , \tag{81}$$

we obtain the system of equations and boundary conditions for $\{u_n^{\pm}\}$ as

$$\sum_{n=0}^{\infty} M_{mn} \frac{\mathrm{d}u_n^+}{\mathrm{d}Z} = -u_m^+ + \sum_{n=0}^{\infty} \left[H_{mn}^{(1)} u_n^+ + H_{mn}^{(2)} u_n^- \right] , \tag{82a}$$

$$-\sum_{n=0}^{\infty} M_{mn} \frac{\mathrm{d}u_n^-}{\mathrm{d}Z} = -u_m^- + \sum_{n=0}^{\infty} \left[H_{mn}^{(1)} u_n^- + H_{mn}^{(2)} u_n^+ \right] , \tag{82b}$$

$$u_m^+|_{Z=0} = \int_0^1 \tilde{P}_m(\mu) b(\mu) \mathrm{d}\mu, \quad m = 0, 1, \dots . \tag{82c}$$

The matrices M, $H^{(1)}$, and $H^{(2)}$ are given by

$$M_{mn} = \int_0^1 \mu \tilde{P}_m(\mu) \tilde{P}_n(\mu) \mathrm{d}\mu, \tag{83a}$$

$$H_{mn}^{(1)} = \frac{1}{2} \int_0^1 \tilde{P}_m(\mu) \int_0^1 h(\mu, \mu') \tilde{P}_n(\mu) \mathrm{d}\mu , \tag{83b}$$

$$H_{mn}^{(2)} = \frac{1}{2} \int_0^1 \tilde{P}_m(\mu) \int_0^1 h(\mu, -\mu') \tilde{P}_n(\mu) \mathrm{d}\mu . \tag{83c}$$

System (82) is to be solved for each $m = 0, 1, \dots$. When we truncate this infinite system at $m = N$, we obtain a finite dimensional system suitable for numerical computations. This truncation yields the double-spherical harmonics of order N approximation,

$$\psi_{\pm}(Z, \mu) \approx \sum_{n=0}^{N} u_n^{\pm}(Z) \tilde{P}_n(\mu), \tag{84}$$

which we denote by DP_N. The objective is to determine the $N + 1$ coefficients $u_n^{\pm}(Z)$ for $n = 0, 1, \dots, N$.

Two useful properties of the half-range Legendre polynomials (79) assist in solving this problem. First, the matrices M, $H^{(1)}$, and $H^{(2)}$ are symmetric. Second, from the theory of generalized Fourier series, it follows that (84), in which $u_n^{\pm}$ for

$n = 0, 1, \ldots, N$ satisfy (82), gives the optimal (in the least-squares sense) polynomial approximation of degree N for $\psi_\pm$. In fact, the advantage of using the half-range specific intensities, $\psi_\pm$, over the full-range specific intensity, ψ, is that former are continuous functions everywhere in their domain of definition. This ensures faster convergence for $\psi_\pm$ as the number of modes N increases since only smoothness over the half-range is required.

Let

$$U^\pm(Z) = (u_0^\pm, u_1^\pm, \ldots, u_N^\pm) \tag{85}$$

denote the vectors of the coefficients corresponding to (80). Using this notation, the problem (82) can be written in matrix notation as

$$\begin{bmatrix} M & 0 \\ 0 & -M \end{bmatrix} \frac{d}{dZ} \begin{bmatrix} U^+ \\ U^- \end{bmatrix} + \begin{bmatrix} \mathbb{I} - H^{(1)} & -H^{(2)} \\ -H^{(2)} & \mathbb{I} - H^{(1)} \end{bmatrix} \begin{bmatrix} U^+ \\ U^- \end{bmatrix} = 0, \quad \text{in } Z > 0, \tag{86a}$$

$$U^+ = b \quad \text{on } Z = 0. \tag{86b}$$

where $\mathbb{I}$ is the $(N + 1) \times (N + 1)$ identity matrix and the entries of the vector b are defined according to

$$b_n = \int_0^1 \tilde{P}_n(\mu) b(\mu) d\mu, \quad n = 0, 1, \ldots, N. \tag{87}$$

We call (86) the DP_N problem for the half space. This is a $2(N + 1) \times 2(N + 1)$ system of equations along with boundary conditions. In what follows, we analyze this system.

5.1 The Generalized Eigenvalue Problem for the DP_N System

We seek a solution of (86) in the form

$$\begin{bmatrix} U^+ \\ U^- \end{bmatrix} = e^{\lambda Z} \begin{bmatrix} V^+ \\ V^- \end{bmatrix}, \tag{88}$$

where λ is an eigenvalue and $[V^+; V^-]$ its associated eigenvector, where, for convenience, we introduce the notation $[U; V] = [U^T, V^T]^T$, where the superscript T denotes the transpose. Substituting (88) into (86a) leads to the $2(N + 1) \times 2(N + 1)$ generalized eigenvalue problem

$$\lambda \begin{bmatrix} M & 0 \\ 0 & -M \end{bmatrix} \begin{bmatrix} V^+ \\ V^- \end{bmatrix} + \begin{bmatrix} \mathbb{I} - H^{(1)} & -H^{(2)} \\ -H^{(2)} & \mathbb{I} - H^{(1)} \end{bmatrix} \begin{bmatrix} V^+ \\ V^- \end{bmatrix} = 0. \tag{89}$$

We establish several important results regarding this problem.

PROPERTY 1. (Symmetry) If λ is an eigenvalue of (89) with eigenvector $[\,V^+;\,V^-\,]$, then $-\lambda$ is an eigenvalue with eigenvector $[\,V^-;\,V^+\,]$.

This property follows from the invariance of (89) under the transformations: $\lambda \rightarrow -\lambda$, $V^+ \rightarrow V^-$, and $V^- \rightarrow V^+$.

PROPERTY 2. (Orthogonality) If λ is an eigenvalue of (89) with eigenvector $[\,V^+;\,V^-\,]$ and ν is an eigenvalue with eigenvector $[\,W^+;\,W^-\,]$, they satisfy the orthogonality relation

$$(\lambda - \nu) \begin{bmatrix} W^+ \\ W^- \end{bmatrix}^T \begin{bmatrix} M & 0 \\ 0 & -M \end{bmatrix} \begin{bmatrix} V^+ \\ V^- \end{bmatrix} = 0. \tag{90}$$

This property follows from left-multiplying (89) for the eigen-triple $(\lambda, V^{\pm})$ by $[\,W^+;\,W^-\,]^T$; left-multiplying (89) for the eigen-triple $(\nu, W^{\pm})$ by $[\,V^+;\,V^-\,]$, and taking the difference between the resulting equations. The terms involving $H^{(1)}$ and $H^{(2)}$ vanish, because these matrices are symmetric. A corollary of PROPERTY 2 is that eigenvectors corresponding to different eigenvalues are M-orthogonal to each other in the sense that

$$\begin{bmatrix} W^+ \\ W^- \end{bmatrix}^T \begin{bmatrix} M & 0 \\ 0 & -M \end{bmatrix} \begin{bmatrix} V^+ \\ V^- \end{bmatrix} = 0. \tag{91}$$

PROPERTY 3. (Nullspace) Zero is an eigenvalue with multiplicity two and only one (proper) corresponding eigenvector. The nullspace also has a one-dimensional generalized eigensolution. In terms of the full-range radiance, the general solution associated with the nullspace is

$$\psi = a_0 + b_0 \left(Z - \frac{\mu}{1-g} \right), \tag{92}$$

where a_0 and b_0 are scalar constants. In terms of the half-range vectors, the generalized eigenvector solution corresponding to (92) is

$$\begin{bmatrix} V^+ \\ V^- \end{bmatrix} = a_0 \begin{bmatrix} \hat{e}_1 \\ \hat{e}_1 \end{bmatrix} + b_0 \left(Z \begin{bmatrix} \hat{e}_1 \\ \hat{e}_1 \end{bmatrix} - \frac{1}{1-g} \begin{bmatrix} m_1 \\ -m_1 \end{bmatrix} \right), \tag{93}$$

where $\hat{e}_1 = (1, 0, \ldots, 0)$ is the unit vector in the "direction" associated with the isotropic half-range Legendre polynomial $\tilde{P}_0 = 1$ and

$$m_1 = M \hat{e}_1 = (M_{00}, M_{01}, \ldots, M_{0N}) \tag{94}$$

is the first column of M, whose entries correspond to the expansion of the coefficients of the functions $\psi_{\pm} = \pm \mu$ in the half-range Legendre polynomials.

This property can be understood as follows. Recall that L has a zero eigenvalue with an associated isotropic eigenfunction i.e., $L P_0(\mu) = 0$. In terms of the half-range

intensities, this isotropic eigenfunction corresponds to $\psi_+ = \psi_- = a_0 \tilde{P}_0(\mu)$ with a_0 denoting a scalar constant. It follows that the vectors of expansion coefficients are given by $U^+ = U^- = a_0 \hat{e}_1$. Hence, the constant a_0 in (92) is in the nullspace. By inspection, the second term in (92) is also a solution of this problem. This implies that the second term in (93) is a generalized eigensolution in the nullspace. Indeed, by PROPERTY 1, the zero eigenvalue has multiplicity two.

5.2   Solving the DP_N Problem

We proceed to formally solve the DP_N problem (86). Recall that (93) is the nullspace solution of the DP_N eigenvalue problem (89) and that all the other eigenvalues come in opposite-sign pairs, $\lambda_{-j} = -\lambda_j$, for $j = 1, 2, \ldots N$, with associated eigenvectors $[V_j^+; V_j^-]$ and $[V_j^-; V_j^+]$, respectively. The general solution of (86a) is

$$\begin{bmatrix} U^+ \\ U^- \end{bmatrix} = a_0 \begin{bmatrix} \hat{e}_1 \\ \hat{e}_1 \end{bmatrix} + b_0 \left(Z \begin{bmatrix} \hat{e}_1 \\ \hat{e}_1 \end{bmatrix} - \frac{1}{1-g} \begin{bmatrix} m_1 \\ -m_1 \end{bmatrix} \right)$$
$$+ \sum_{j=1}^{N-1} \left(\begin{bmatrix} V_j^- \\ V_j^+ \end{bmatrix} e^{-\lambda_j Z} a_j + \begin{bmatrix} V_j^+ \\ V_j^- \end{bmatrix} e^{\lambda_j Z} b_j \right) , \qquad (95)$$

where a_j, and b_j for $j = 0, 1, \ldots, N$ are coefficients to be determined. Requiring that (95) remain bounded for all $Z > 0$ leads to setting $b_0 = b_1 = \ldots = b_N = 0$. Therefore,

$$\begin{bmatrix} U^+ \\ U^- \end{bmatrix} = a_0 \begin{bmatrix} \hat{e}_1 \\ \hat{e}_1 \end{bmatrix} + \sum_{j=1}^{N} \begin{bmatrix} V_j^- \\ V_j^+ \end{bmatrix} e^{-\lambda_j Z} a_j. \qquad (96)$$

Requiring that (96) satisfy the boundary condition (86b), we obtain

$$a_0 \hat{e}_1 + \sum_{j=1}^{N} V_j^- a_j = b. \qquad (97)$$

Equation (97) is a linear system for the coefficients a_j, $j = 0, 1, \ldots, N$, which can be readily solved. The solution of (72) is then obtained from (96) and the DP_N expansion (80).

We can formally write the solution of (97) as $a = \Gamma I_0$, where a denotes the vector of expansion coefficients and Γ denotes the inverse of the matrix whose first column is $\hat{e}_1$ and whose subsequent columns are V_j^- for $j = 1, \ldots, N$, i.e.,

$$\Gamma = \left[\hat{e}_1; V_1^-; \cdots; V_N^- \right]^{-1} . \qquad (98)$$

The matrix Γ gives the mapping from the boundary data to the coefficients making up the solution of (97). Let $\boldsymbol{g}_0^T$ denote the first row of Γ. It follows that

$$a_0 = \boldsymbol{g}_0^T \boldsymbol{b}. \tag{99}$$

Therefore, the inner product of $\boldsymbol{g}_0$ with the boundary data is the mapping from the boundary data to the only non-decaying mode of (96). This operation is precisely the DP_N approximation of the operator $\mathscr{G}_s^0$ introduced in Sect. 4.3.

For boundary layer solutions about $z = 1$, we need to solve

$$-\mu \frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}\tilde{Z}} + L\tilde{\psi} = 0 \quad \text{in} \quad \tilde{Z} > 0, \tag{100}$$

subject to

$$\tilde{\psi}|_{Z=0} = \tilde{b} \quad \text{on} \ -1 \le \mu < 0. \tag{101}$$

By replacing μ with $-\mu$, we obtain the same problem as (72) subject to boundary condition (73). It follows that DP_N solution is

$$\begin{bmatrix} \tilde{U}^+ \\ \tilde{U}^- \end{bmatrix} = \tilde{a}_0 \begin{bmatrix} \hat{e}_1 \\ \hat{e}_1 \end{bmatrix} + \sum_{j=1}^{N-1} \begin{bmatrix} V_j^+ \\ V_j^- \end{bmatrix} e^{-\lambda_j Z} \tilde{a}_j, \tag{102}$$

with the vector of coefficients, $\tilde{a} = (a_0, a_1, \ldots, a_N)$ defined as $\tilde{a} = \Gamma \tilde{b}$ with

$$\tilde{b}_n = \int_0^1 \tilde{P}_n(\mu)\tilde{b}(-\mu)\mathrm{d}\mu, \quad n = 0, 1, \ldots, N. \tag{103}$$

Note that V_j^+ and V_j^- are switched in (102) from what appears in (96) due to the replacement of μ by $-\mu$.

6 Procedure for Computing the Asymptotic Approximation of Boundary Layer Theory

We now have all of the components needed to evaluate the uniformly valid asymptotic approximation of the RTE in the strong-scattering regime. In what follows, we summarize the results from above in a procedure to evaluate the asymptotic approximation valid to $\mathscr{O}(\varepsilon^2)$.

To compute the asymptotic approximation of the solution to

$$\varepsilon\mu \frac{\mathrm{d}I}{\mathrm{d}z} + \varepsilon^2 I + LI = 0 \quad \text{in} \ 0 < z < 1, \tag{104}$$

for a specified redistribution function subject to boundary conditions

$$I|_{z=0} = b \quad \text{on} \quad 0 < \mu \leq 1, \tag{105a}$$

$$I|_{z=1} = 0 \quad \text{on} \quad -1 \leq \mu < 0, \tag{105b}$$

we follow the procedure given below.

1. Interior solution

   a. For the DP_N method, set N and compute (83a), (83b), (83c), (87), and (94) using Gauss quadrature to obtain M, $H^{(1)}$, $H^{(2)}$, b, and m_1, respectively.
   b. Construct and solve generalized eigenvalue problem (89) for the eigenvalues $\lambda_{\pm j}$ and eigenvectors $V_j^{\pm}$, for $j = 0, 1, 2, \ldots N$.
   c. Use the solution from Step 1(b) to compute the matrix Γ given in (98) and set g_0^T to be its first row.
   d. Compute the coefficients

$$\beta_0 = g_0^T \hat{e}_1, \tag{106a}$$

$$\beta_1 = g_0^T m_1, \tag{106b}$$

$$\bar{b} = g_0^T b. \tag{106c}$$

   e. Using the results from Step 1(d), solve the diffusion approximation given by (71) to obtain $\rho(z)$.
   f. Compute the half-range interior solutions,

$$\Phi_{\pm}(z, \mu) = \rho(z) \mp \varepsilon \mu \rho'(z)/(1 - g), \tag{107}$$

   for $z \in (0, 1)$, and $\mu \in (0, 1]$.

2. Boundary layer solutions

   a. Compute the coefficients

$$a = \Gamma[b - \rho(0)\hat{e}_1 + \varepsilon m_1 \rho'(0)/(1 - g)], \tag{108}$$

$$\tilde{a} = \Gamma[-\rho(1)\hat{e}_1 - \varepsilon m_1 \rho'(1)/(1 - g)]. \tag{109}$$

   b. Using the results from Step 2(a), compute

$$\begin{bmatrix} U^+ \\ U^- \end{bmatrix} = \sum_{j=1}^{N} \begin{bmatrix} V_j^- \\ V_j^+ \end{bmatrix} e^{-\lambda_j z/\varepsilon} a_j, \tag{110}$$

$$\begin{bmatrix} \tilde{U}^+ \\ \tilde{U}^- \end{bmatrix} = \sum_{j=1}^{N} \begin{bmatrix} V_j^+ \\ V_j^- \end{bmatrix} e^{-\lambda_j(1-z)/\varepsilon} \tilde{a}_j, \tag{111}$$

c. Using the results from Step 2(b), compute the half-range boundary layer solutions

$$\psi_\pm^{(1)}(z, \mu) = \sum_{n=0}^{N} u_n^\pm(z) \tilde{P}_n(\mu) , \tag{112}$$

$$\psi_\pm^{(2)}(z, \mu) = \sum_{n=0}^{N} \tilde{u}_n^\pm(z) \tilde{P}_n(\mu) . \tag{113}$$

3. Asymptotic approximation

Compute the asymptotic approximation of the half-range specific intensities,

$$I_\pm(z, \mu) = \Phi_\pm(z, \mu) + \psi_\pm^{(1)}(z, \mu) + \psi_\pm^{(2)}(z, \mu) , \tag{114}$$

for $z \in [0, 1]$, and $\mu \in (0, 1]$.

We call (114) the uniformly valid asymptotic approximation of the boundary layer theory to $\mathcal{O}(\varepsilon^2)$. It gives an approximation to the solution of the RTE in the entire domain with an error that scales as $\mathcal{O}(\varepsilon^2)$. The interior solution $\Phi_\pm$ is derived from the diffusion approximation. The boundary layer solutions $\psi_\pm^{(1)}$ and $\psi_\pm^{(2)}$ give corrections to the diffusion approximation near boundaries $z = 0$ and $z = 1$, respectively.

7 Deriving Two-Flux Theory

Using the results from boundary layer theory discussed above, we derive two-flux theory. To do so, we introduce the half-range fluxes,

$$F_\pm(z) = \int_0^1 I_\pm(z, \mu)\mu \, \mathrm{d}\mu . \tag{115}$$

Multiplying μ to (107) and integrating with respect to μ, we find that the associated fluxes are

$$F_\pm(z) = \frac{1}{2}\rho(z) \mp \frac{\varepsilon}{3(1 - g)}\rho'(z) . \tag{116}$$

Equation (116) establishes a relationship between $F_\pm$ and the diffusion approximation.

Adding and subtracting the two equations in (116) for $F_\pm$, we get

$$F_+(z) + F_-(z) = \rho(z) , \tag{117a}$$

$$F_+(z) - F_-(z) = -\frac{2\varepsilon}{3(1 - g)}\rho'(z) . \tag{117b}$$

Differentiating (117) with respect to z gives

$$F'_+(z) + F'_-(z) = \rho'(z) , \tag{118a}$$

$$F'_+(z) - F'_-(z) = -\frac{2\varepsilon}{3(1-g)}\rho''(z) . \tag{118b}$$

Solving for ρ' in (117b) and substituting that result into (118a), we obtain

$$F'_+(z) + F'_-(z) = -\frac{3(1-g)}{2\varepsilon}\left[F_+(z) - F_-(z)\right] . \tag{119}$$

Substituting (71a) into the right-hand side of (118b), we obtain

$$F'_+(z) - F'_-(z) = -2\varepsilon\alpha\rho(z) . \tag{120}$$

Solving for ρ from (117a) and substituting into (120), we get

$$F'_+(z) - F'_-(z) = -2\varepsilon\alpha[F_+(z) + F_-(z)] . \tag{121}$$

By adding and then subtracting (119) and (121), dividing by 2, and rearranging terms, we arrive at

$$F'_+(z) = -\left[\frac{3(1-g)}{4\varepsilon} + \varepsilon\alpha\right]F_+ + \left[\frac{3(1-g)}{4\varepsilon} - \varepsilon\alpha\right]F_- , \tag{122a}$$

$$F'_-(z) = \left[\frac{3(1-g)}{4\varepsilon} + \varepsilon\alpha\right]F_- - \left[\frac{3(1-g)}{4\varepsilon} - \varepsilon\alpha\right]F_+ . \tag{122b}$$

System (122) can be written as

$$F'_+(z) = -(\tilde{K} + \tilde{S})F_+ + \tilde{S}F_- , \tag{123a}$$

$$F'_-(z) = (\tilde{K} + \tilde{S})F_+ - \tilde{S}F_- , \tag{123b}$$

with coefficients

$$\tilde{K} = 2\varepsilon\alpha , \tag{124a}$$

$$\tilde{S} = \frac{3}{4\varepsilon}(1-g) - \varepsilon\alpha . \tag{124b}$$

The equations in (123) are the two-flux equations, but written with respect to scaling given in (20). We can write them in terms of the optical depth τ, defined in (9), and use the albedo (10) to express the two-flux system as

$$F'_+(\tau) = -(\tilde{K} + \tilde{S})F_+(\tau) + \tilde{S}F_-(\tau) , \tag{125a}$$

$$F'_-(\tau) = (\tilde{K} + \tilde{S})F_+(\tau) - \tilde{S}F_-(\tau) , \tag{125b}$$

with

$$\tilde{K} = 2(1 - \varpi_0), \tag{126a}$$

$$\tilde{S} = \frac{3}{4}(1 - g) - (1 - \varpi_0) . \tag{126b}$$

Alternatively, by transforming back to the dimensional z variable using (9) and (10), we arrive at the familiar two-flux equations in dimensional form,

$$F'_+(z) = -(K + S)F_+(z) + SF_-(z), \tag{127a}$$

$$F'_-(z) = (K + S)F_+(z) - SF_-(z) , \tag{127b}$$

with

$$K = 2\kappa_a, \tag{128a}$$

$$S = \frac{3}{4}\kappa_s(1 - g) - \frac{1}{4}\kappa_a(1 + 3g) . \tag{128b}$$

These results are the same as derived by Sandoval and Kim (2014).

We have derived the two-flux equations (127) above from computing the half-range fluxes associated with the interior solution in boundary layer theory. This derivation shows that in a strongly scattering medium, two-flux theory is entirely consistent with the diffusion approximation. In fact, it is just a reformulation of the diffusion approximation. Consequently, two-flux theory is asymptotically accurate in the interior of the medium.

7.1 Limitations of Two-Flux Theory

Two-flux theory provides a simple and intuitive picture of light propagation in a scattering medium. For that reason, it is very useful. However, it is important to point out its limitations which we list below.

LIMITATION I. Two-flux theory is only valid for one-spatial dimension for problems with azimuthal symmetry.
LIMITATION II. Two-flux theory is only accurate for high-albedo media.
LIMITATION III. Two-flux theory is only valid in the interior of the domain. This theory does not accurately determine the fluxes leaving the domain.

LIMITATION I is inherent in the assumption of azimuthal symmetry leading to the one-dimensional RTE (4). We discuss a possible extension to three-dimensional space in Sect. 9. LIMITATION II and LIMITATION III are manifest from our derivation of System (127) from the interior solution (107) given in terms of the solution of the diffusion approximation (71). Boundary layer theory provides an asymptotic solution in the limit as $\varepsilon \to 0^+$ corresponding to $\varpi_0 \to 1$. Furthermore, we have learned from

boundary layer theory that the interior solution requires corrections in the form of boundary layer solutions to accurately capture the intensity near the boundary.

7.2 Boundary Conditions

The two-flux equations need to be supplemented with boundary conditions. These boundary conditions are usually determined by computing the half-range fluxes from boundary conditions of the RTE in the following way. Multiplying μ to (8a) and (8b) integrating those results with respect to μ, we find that

$$F_+(0) = \int_0^1 b(\mu)\mu \, \mathrm{d}\mu, \tag{129a}$$

$$F_-(1) = 0. \tag{129b}$$

These boundary conditions intuitively prescribe the fluxes incident on the boundaries $z = 0$ and $z = 1$. However, if we consider the fact that the two-flux equations are derived directly from the interior solution given by the diffusion approximation, these boundary conditions may not be appropriate. In boundary layer theory, this interior solution requires the addition of boundary layer solutions to correct it near the boundaries. We call the solution of the two-flux equations given by (123) with boundary conditions (129) *naïve two-flux theory*.

As an alternative to boundary conditions (129) in naïve two-flux theory, we derive boundary conditions consistent with those given in (71b) from boundary layer theory. Substituting (117) into (71b), we find that

$$F_+(0) = \frac{1}{\beta_0 + \frac{3}{2}\beta_1}\bar{b} - \frac{\beta_0 - \frac{3}{2}\beta_1}{\beta_0 + \frac{3}{2}\beta_1}F_-(0), \tag{130a}$$

$$F_-(1) = \frac{\beta_0 - \frac{3}{2}\beta_1}{\beta_0 + \frac{3}{2}\beta_1}F_+(1). \tag{130b}$$

In contrast to boundary conditions (129), these boundary conditions include an effective reflection at the boundary with reflection coefficient $R = (\beta_0 - 3\beta_1/2)/(\beta_0 + 3\beta_1/2)$. Since these boundary conditions are derived from the asymptotic theory, we call the solution of the two-flux equations given by (123) with boundary conditions (130) *asymptotic two-flux theory*.

We expect a large difference in the accuracy between the naïve and asymptotic two-flux theories. The entire solution of the two-flux equations depends on the boundary conditions. Hence, differences in boundary conditions lead to differences in the entire solution. Moreover, we identified in LIMITATION III that two-flux theory cannot accurately capture the fluxes leaving the domain. Naïve two-flux theory insists that the boundary conditions are the incident fluxes on the boundaries. In contrast,

asymptotic two-flux theory anticipates the need for boundary layer solutions to overcome LIMITATION III. We investigate the differences between these different two-flux theories in the computational results.

8 Computational Results

In this section, we compare the radiative transfer theory and the approximations discussed above. We seek to solve the RTE (21) with $\alpha = 1$ for different values of ε (or the albedo), and the Henyey-Greenstein redistribution function, defined as (Van de Hulst 2012)

$$h_{\text{HG}}(\mu, \mu'; g) = \frac{(1 - g^2)E(k)}{\pi |\beta - \gamma|\sqrt{|\beta + \gamma|}} , \tag{131a}$$

$$\beta = 1 - g^2 - 2g\mu\mu' \tag{131b}$$

$$\gamma = 2g\sqrt{(1 - \mu^2)(1 - \mu'^2)} , \tag{131c}$$

where, in (131a), $E(k)$ is the complete elliptic integral of the second kind, defined as

$$E(k) = \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2 \theta} \, d\theta , \tag{132}$$

evaluated with $k = 2\gamma/|\beta + \gamma|$. The redistribution function (131) is obtained from the three-dimensional Henyey-Greenstein scattering phase function, defined as

$$p_{\text{HG}}(\boldsymbol{\Omega}, \boldsymbol{\Omega}'; g) = \frac{1}{4\pi} \frac{1 - g^2}{\left(1 + g^2 - 2g\,\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}'\right)^{3/2}} , \tag{133}$$

using the integral defined in (6).

In particular, we consider (21) in the domain $0 < z < 1$ with the boundary conditions

$$I|_{z=0} = \mu^2, \text{ on } 0 < \mu \leq 1 , \tag{134a}$$

$$I|_{z=1} = 0 , \text{ on } -1 \leq \mu < 0 . \tag{134b}$$

Upon solution of this problem, we compute the half-range fluxes

$$F_{\pm}(z) = \int_0^1 I_{\pm}(z, \mu)\mu d\mu. \tag{135}$$

8.1 Comparison of Approximations

We compute the approximations of the problem (21) with boundary conditions (134) using the following three theories.

- Naïve two-flux theory [(123) with boundary conditions (129)];
- Asymptotic two-flux theory [(123) with boundary conditions (130)];
- Boundary layer theory [Procedure given in Sect. 6].

In particular, we study $F_{\pm}(z)$ computed from all of these theories over $0 \leq z \leq 1$.

To begin with, we solve this problem with $\varepsilon = 0.01$, which corresponds to $\varpi_0 \approx 0.9999$ [see (23)], and $g = 0.4$, which corresponds to a moderately forward-scattering medium. We solve this problem using the DP_N method with $N = 12$. We use 501 grid points in z, and evaluate the integrals in (131) using the 128 point Gauss-Legendre quadrature rule. In what follows, we denote the solution thus obtained the DP_{12} solution. To validate using this method, we compared results from the DP_{12} solution to those from a highly resolved computation using the discrete ordinate method with the double-Gauss quadrature rule (see Thomas and Stamnes 2002, for example). This method has been shown to be pointwise convergent (Keller 1960). Through those comparisons, we have found that for this problem, the DP_{12} solution has a maximum relative error that is less than 10^{-6}. For this reason, we consider the DP_{12} solution as the benchmark, to which all other approximations evaluated here are compared.

Let $F_{\pm}(z)$ denote to the benchmark (DP_{12}) solutions and let $\tilde{F}_{\pm}(z)$ denote the approximate solutions. In what follows we refer to the absolute maximum errors, the relative maximum errors, and the root-mean square (RMS) errors of the solutions. These errors are defined as

$$\text{absolute maximum error} = \max_{z} \left\{ |F_+(z) - \tilde{F}_+(z)| \, , \, |F_-(z) - \tilde{F}_-(z)| \right\},$$

$$\text{relative maximum error} = \max_{z} \left\{ \frac{|F_+(z) - \tilde{F}_+(z)|}{|F_+(z)|} \, , \, \frac{|F_-(z) - \tilde{F}_-(z)|}{|F_-(z)|} \right\},$$

$$\text{RMS error} = \max \left\{ \sqrt{\frac{1}{N_z} \sum_{i=1}^{N_z} |F_+(z_i) - \tilde{F}_+(z_i)|^2} \, , \, \sqrt{\frac{1}{N_z} \sum_{i=1}^{N_z} |F_-(z_i) - \tilde{F}_-(z_i)|^2} \right\},$$

respectively, where $\{z_i\}_{i=1}^{N_z}$ denote the discrete points along the z direction.

Figure 1 shows a comparison between the DP_{12}, naïve two-flux, and asymptotic two-flux solutions. The top row of plots show these solutions for $0 \leq z \leq 1$. The bottom row of plots show these solutions near the boundary at $z = 0$. The DP_{12} solution changes rapidly near the boundary at $z = 0$ corresponding to the boundary layer. Because the solution itself is very small near $z = 1$, the rapid change near the boundary $z = 1$ is not apparent in Fig. 1. Regardless, neither the naïve nor the asymptotic two-flux solutions accurately capture these rapid changes near the boundary. The naïve two-flux solution is accurate at the boundaries, but not away from them. The absolute error made by the the naïve two-flux solution for $F_+(1)$ is 4.224×10^{-4}

and for $F_-(0)$ is 8.689×10^{-4}. Even though it captures the qualitative behavior of $F_\pm(z)$ over $0 < z < 1$, it exhibits a large quantitative error. In contrast, the asymptotic two-flux solution accurately approximates the DP_{12} solution for $0 < z < 1$, but makes a substantial error at the boundary layer near $z = 0$.

The relative maximum error of the naïve two-flux solution is 0.0930. This maximum error is attained in the interior of the domain. The relative maximum error of the asymptotic two-flux solution is 0.0988. This maximum error is attained at $z = 0$. From the results in Fig. 1, we observe that the asymptotic two-flux theory is accurate over a much larger portion of the domain than the naïve two-flux theory. This difference is captured better by the RMS error. The RMS error of the naïve two-flux solution is 0.0136, whereas, the RMS error of the asymptotic two-flux theory is 0.00163. With respect to the the RMS error, the asymptotic two-flux theory is an order of magnitude more accurate than the naïve two-flux theory.

Figure 2 shows a comparison between the half-range interior solutions, $\Phi_\pm$, boundary layer solutions, $\Psi_\pm^{(1)} + \Psi_\pm^{(2)}$, and the uniformly valid asymptotic solution of boundary layer theory, which is the sum given in (114). Here, we see how the boundary layer solutions correct the interior solution near the boundaries. They are exponentially small in the interior of the domain. Consequently, they do not interfere with the approximation made by the interior solution there. Because the boundary layer solutions take care of the boundary conditions and the rapid change in the solution near the boundaries, they free the interior solution from having to satisfy boundary conditions it is not capable of satisfying which, in turn, allows it to accurately capture the solution in the interior of the domain. It is in this way that
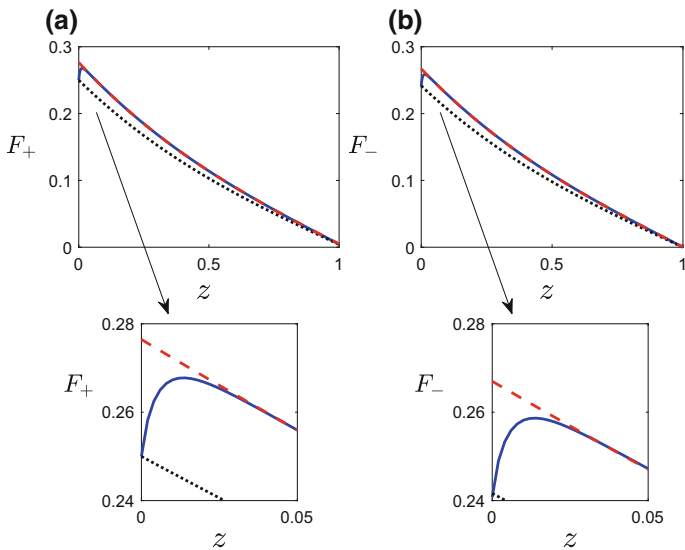
Fig. 1 Comparison between the DP_{12} solution (solid), the naïve two-flux solution (dots), and asymptotic two-flux solution (dashes). Here $\varepsilon = 0.01$ ($\varpi_0 \approx 0.9999$) and $g = 0.4$. Panels (a) and (b) show F_+ and F_-, respectively. The bottom row of panels "zoom in" on the solutions inside the boundary layer near $z = 0$
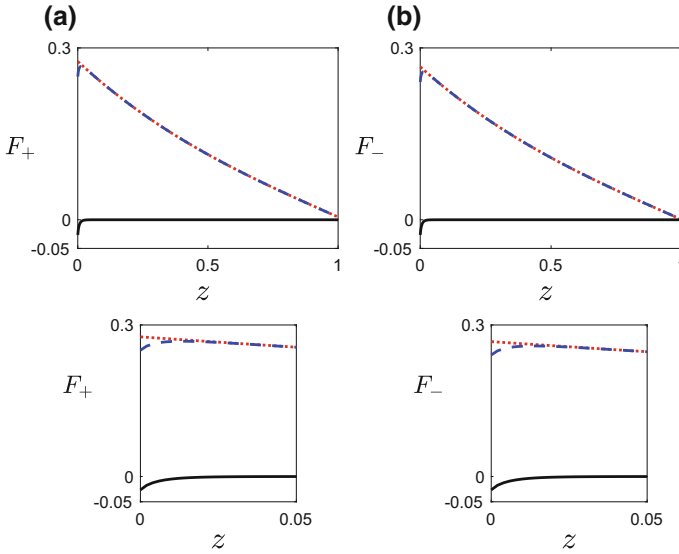
Fig. 2 Comparison between the interior solution (dots), the boundary layer solution (solid), and uniformly valid asymptotic solution (dashes) of boundary layer theory, for the same problem as in Fig. 1. The bottom row of panels "zoom in" on the solutions inside the boundary layer near $z = 0$

boundary layer theory provides a uniformly accurate approximation throughout the entire domain.

Figure 3 shows a comparison between the solutions computed using DP_{12}, the uniformly valid asymptotic solution of boundary layer theory, and asymptotic two-flux theory. Since asymptotic two-flux theory is consistent with the diffusion approximation, it is equivalent to the interior solution in boundary layer theory. The interior solution does not satisfy the boundary conditions prescribed with the RTE, so it makes a significant error near the boundary $z = 0$. On the other hand, because boundary layer theory corrects the diffusion approximation by including boundary layer solutions near each of the boundaries, it provides a uniformly valid asymptotic approximation for all $0 \leq z \leq 1$. In particular, these boundary layer solutions accurately capture the rapid change in the solution near $z = 0$ and $z = 1$. In particular, the relative maximum error of the uniformly valid asymptotic approximation is approximately 8.343×10^{-4} with an RMS error of 1.135×10^{-4}. As a result, the uniformly valid asymptotic approximation and the DP_{12} solution are indistinguishable to the eye. These results clearly show that boundary layer theory provides a very accurate approximation over the whole domain. It is vastly superior over both of the two-flux theories, as well as the diffusion approximation, since it includes the boundary layer solutions that take into account the rapid changes near the boundaries.

The derivation of two-flux theory indicates that it is valid only in the interior of the domain of a strongly scattering medium. The results shown here indicate that asymptotic two-flux theory is more accurate than the naïve two-flux theory everywhere except near the boundaries. If one only considers the fluxes exiting the
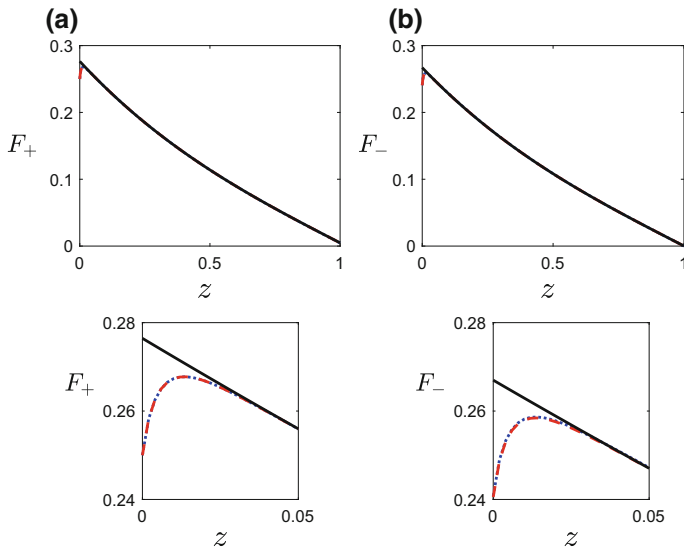
Fig. 3 Comparison between the DP_{12} solution (dots), the uniformly valid asymptotic approxima-
tion (dashes), and the asymptotic two-flux theory (solid) for the same problem as in Fig. 1. The
bottom row of panels "zoom in" on the solutions inside the boundary layer near $z = 0$

domain, it appears that the naïve two-flux theory provides a more accurate approxi-
mation. However, using the naïve two-flux theory in this way is problematic because
it achieves this approximation at the expense of a substantial error everywhere else.
In contrast, boundary layer theory provides a uniformly accurate approximation over
the entire domain. Moreover, it is substantially more accurate.

8.2 Dependence on the Albedo

We now consider the accuracy of these approximations for different values of ε,
which is directly related to the albedo through (23). In particular, we compute the
relative errors of the different approximations with $g = 0.4$ for $0.001 \leq \varepsilon \leq 0.25$
corresponding to $0.94 < \varpi_0 < 1$. According to (23), $1 - \varpi_0 = \mathcal{O}(\varepsilon^2)$ and so is the
error of the asymptotic theory. The results in Fig. 4 show that the relative errors of
both the naïve and asymptotic two-flux theories are approximately 10%. In contrast,
the error of the uniformly valid asymptotic solution of boundary layer theory scales
as $\mathcal{O}(1 - \varpi_0)$ as predicted by the asymptotic theory.

Fig. 4 Log-log plot of the errors of the naïve two-flux solution (dots), the asymptotic two-flux solution (dashes), and the uniformly valid asymptotic solution (solid) as functions of $1 - \varpi_0$. All other parameters are the same as in Fig. 1
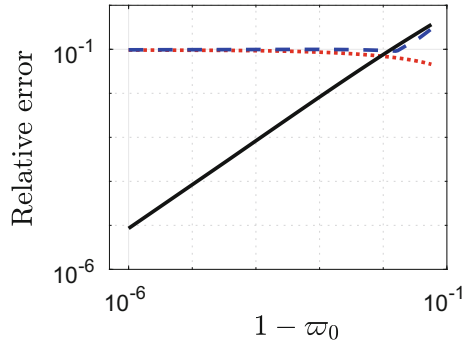
Fig. 5 a Log plot of the errors of the naïve two-flux solution (dots), the asymptotic two-flux solution (dashes), and the uniformly valid asymptotic solution (solid) as functions of the anisotropy parameter g. All other parameters are the same as in Fig. 1
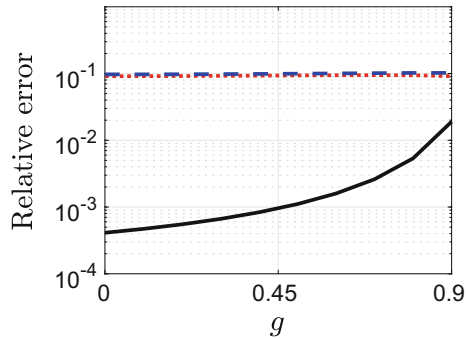
8.3 Dependence on the Anisotropy Factor

The only other parameter in this problem besides the albedo, ϖ_0, is the anisotropy factor, g. The results shown in Sect. 8.1 were for $g = 0.4$ corresponding to mildly anisotropic scattering. We now consider the accuracy of these approximations for different values of g in Eq. (131). According to the uniformly valid asymptotic theory, the error is $\mathcal{O}(\varepsilon^2)$. Thus, the leading order behavior of the error is given by $C\varepsilon^2$, where C is some constant that is independent of ε (and consequently, ϖ_0). However, it may depend on g. In what follows, we seek to determine the dependence of C on g.

Figure 5 shows the relative errors of the naïve and asymptotic two-flux theories, and the uniformly valid asymptotic solution of boundary layer theory. These results show that the errors for both the naïve and asymptotic two-flux theories are approximately 10% over all values of g. In contrast, the error of the uniformly valid asymptotic solution of boundary layer theory is two order of magnitude smaller or less for $0 \leq g \leq 0.9$. However, it appears to increase and become large as $g \to 1$.

8.4 Forward-Peaked Scattering

The results shown in Fig. 5 are for ε held fixed at $\varepsilon = 0.01$. Boundary layer theory ensures that the error of the asymptotic solution will decrease quadratically as $\varepsilon \to 0$ for any fixed value of g. It does not account for sharply forward-peaked scattering corresponding to the asymptotic limit, $g \to 1$. With sharply forward-peaked scattering, the transport mean-free path, $\ell_{tr} = 1/[\kappa_s(1 - g)]$, is much larger than the scattering mean-free path, $\ell_s = 1/\kappa_s$, since $g \approx 1$. In this parameter regime, there is a wide separation between the scattering and diffusion length scales. As a result, anisotropic, sharply forward-peaked scattering inhibits the onset of the diffusion limit. Many more multiple scattering events are required for the radiance to become nearly isotropic compared with the mildly anisotropic scattering case. It follows that the penetration depth for the onset of the diffusion limit is much larger for this case. At intermediate penetration depths, the character of highly anisotropic scattering is different and requires additional considerations. Larsen (1999) gives an asymptotic analysis of this problem leading to useful asymptotic expansions for the scattering operator.

To investigate the case of anisotropic forward-peaked scattering, we show results for $g = 0.9$. In Fig. 6, we compare the DP_{12}, naïve two-flux, and asymptotic two-flux solutions just as we have done in Fig. 1 for the $g = 0.4$ case. The overall qualitative behaviors are the same as for $g = 0.4$. However, we observe here larger errors made by both the naïve and asymptotic two-flux solutions. The errors for F_- are larger than those for F_+. The relative maximum error of the naïve two-flux solution is 0.0913
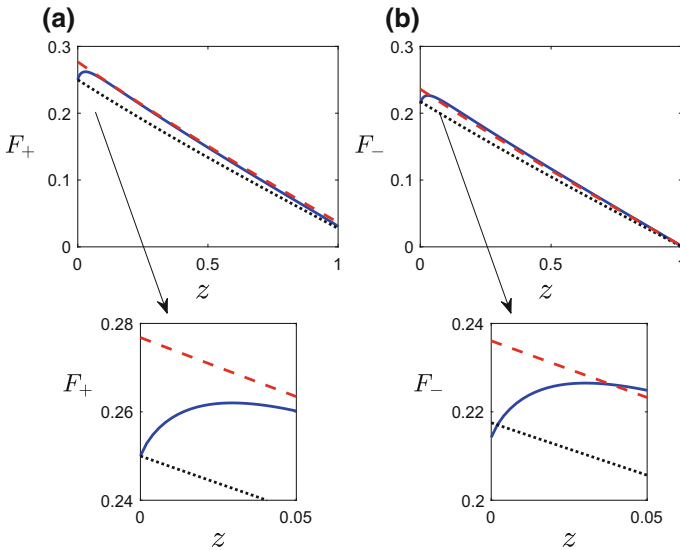
Fig. 6 Same as Fig. 1 for $g = 0.9$

(a)

F_+

(b)

F_-

F_+

F_-
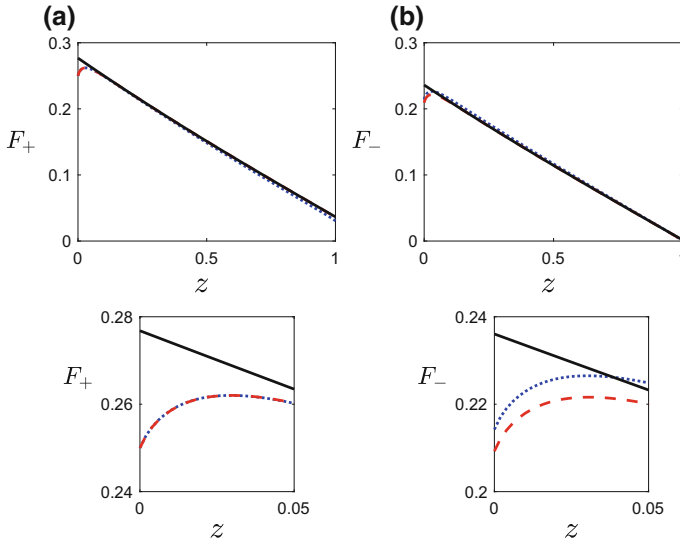
Fig. 7 Same as Fig. 3 with $g = 0.9$

and the relative maximum error of the asymptotic two-flux solution is 0.1022. The RMS errors are 0.0157 and 0.0040 for the naïve and asymptotic two-flux solutions, respectively.

In Fig. 7 we show comparisons with the DP_{12} solution, the uniformly valid asymptotic solution of boundary layer theory, and asymptotic two-flux theory, just as Fig. 2 for the $g = 0.4$ case. Here, we find that these approximations are not as accurate as they were for the $g = 0.4$ case. The aforementioned errors of the asymptotic two-flux solution carry over to the asymptotic solution of boundary layer theory. In particular, the larger errors for F_- affect the boundary layer solution which, in turn, exhibits a noticeable error in the lower right plot of Fig. 6. The relative maximum error for the boundary layer theory asymptotic approximation is 0.0190 and the RMS error is 0.0030.

Just as Fig. 5 indicates, we have found that the performance of all approximations is worse for $g = 0.9$. The behavior of the error of the asymptotic solution of boundary layer theory will decrease quadratically with ε. However, for this chosen value of $\varepsilon = 0.01$ and $g = 0.9$, we find that the error is much larger than the case when $g = 0.4$.

8.5 Four-Flux Theory and the DP_1 Approximation

Two-flux theory is valid only in the interior of the domain. The eigenvalues of (123) are given by

$$\lambda = \pm\sqrt{\tilde{K}(\tilde{K} + 2\tilde{S})} = \pm\sqrt{3\alpha(1 - g)}, \tag{136}$$

where we have used (124). Note that these eigenvalues correspond exactly to those for the diffusion equation given in (71a). This result is not surprising since two-flux theory is derived from the interior solution governed by the diffusion approximation.

An improvement over two-flux theory, called four-flux theory, incorporates an additional length scale in the problem. This length scale is smaller than that given in (136) and, from what we now understand from boundary layer theory, attempts to resolve the rapid changes of the solution near the boundaries. Rather than give a phenomenological method to determine this additional length scale, we consider the DP_N method with $N = 1$, which we call the DP_1 approximation.

If we apply the DP_1 approximation to (21), we obtain a 4×4 system. By PROP-ERTY 1 of generalized eigenvalue problem (89), the DP_1 approximation will produce 2 distinct length scales corresponding to the 2 distinct $\pm$ pairs of eigenvalues. One of those length scales corresponds to (136). The other length scale is the additional length scale in four-flux theory. It is one eigenmode that seeks to approximate the boundary layer solutions. For this reason, we call the four-flux solutions the results from computing the half-range fluxes $F_\pm$ from the solution of the DP_1 system.

Figure 8 shows a comparison of results computed using DP_{12}, the boundary layer solution, and four-flux theory. Here, we see that the four-flux solution captures the rapid changes in the solution near the boundaries better than the two-flux solution does. The maximum relative error made by the four-flux solution is 0.0129 and the
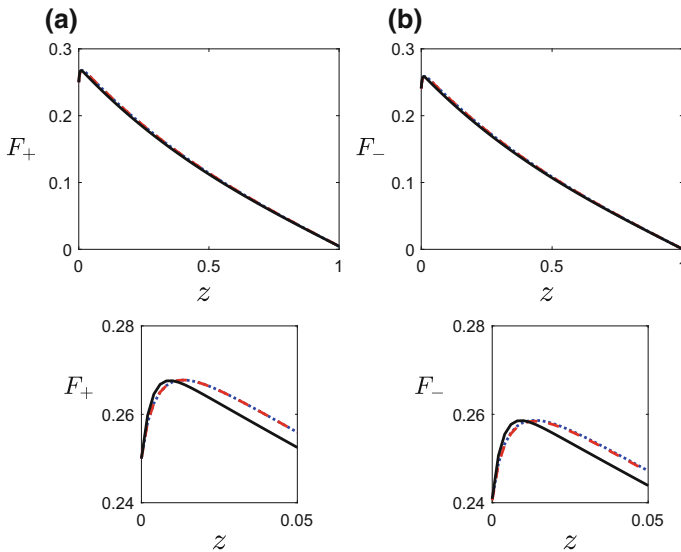
Fig. 8 Comparison between the DP_{12} solution (dotted blue), the uniformly valid asymptotic approximation (dashed red), and four-flux solution (solid black) for the same problem as in Fig. 1. The bottom row of panels "zoom in" on the solutions inside the boundary layer near $z = 0$

RMS error is 0.00185. This maximum relative error is 7 times smaller than those for the naïve and asymptotic two-flux solutions. This RMS error is an order of magnitude smaller than that for the naïve two-flux solution, but of the same order as the RMS error for the asymptotic two-flux solution. However, boundary layer theory provides a vastly better approximation than four-flux theory.

Four-flux theory provides a substantial improvement over two-flux theory through the inclusion of an additional length scale. Rather than seeking to derive this length scale phenomenologically, we consider instead computing the half-range fluxes from the DP_1 approximation. The DP_1 approximation has 2 distinct length scales: one corresponds to the two-flux/diffusion length scale given in (136), and the other attempts to resolve the rapid changes near the boundaries. Although four-flux theory qualitatively captures the rapid changes in the solution near the boundaries, it is less accurate than the asymptotic approximation derived from boundary layer theory.

9 Extension of Two-Flux Theory for Three-Dimensional Radiative Transfer

Many physical problems of interest require solving the three-dimensional RTE. For this reason, it is tempting to seek a two-flux theory for the three-dimensional RTE. Such a theory may be used to study beam propagation and scattering in a plane-parallel medium, for example. However, LIMITATION I in Sect. 7.1 states that two-flux theory is only valid for one-spatial dimension for problems with azimuthal symmetry. Nonetheless, using the connection between two-flux theory and the diffusion approximation established above, we can develop what amounts to a two-flux theory for the three-dimensional RTE.

The diffusion equation in three dimensions is given by

$$D\rho_{zz} + D\nabla_\perp^2 \rho - \alpha\rho = 0 \,, \tag{137}$$

where we have introduced the diffusion coefficient $D = [3(1 - g)]^{-1}$, for convenience, and $\nabla_\perp^2 = \partial_{xx}^2 + \partial_{yy}^2$ denoting the transverse Laplacian. In boundary layer theory for the three-dimensional RTE, the interior solution is given by

$$\Phi = \rho - \varepsilon 3D\boldsymbol{\Omega} \cdot \nabla\rho + \mathcal{O}(\varepsilon^2), \tag{138}$$

The half-range fluxes associated with this interior solution can be calculated as

$$F_\pm(x, y, z) = \int_0^{2\pi} \int_0^1 \Phi_\pm(x, y, z, \mu, \varphi)\mu \mathrm{d}\mu \mathrm{d}\varphi = \frac{1}{2}\rho \mp \varepsilon D\rho_z + \mathcal{O}(\varepsilon^2). \tag{139}$$

Thus, we find that

$$F_+ + F_- = \rho, \tag{140a}$$
$$F_+ - F_- = -\varepsilon 2 D \rho_z. \tag{140b}$$

By taking the partial derivative with respect to z, we obtain

$$F_{+,z} + F_{-,z} = \rho_z = -\frac{1}{2\varepsilon D}(F_+ - F_-), \tag{141a}$$
$$F_{+,z} - F_{-,z} = -\varepsilon 2 D \rho_{zz} = -\varepsilon 2(\alpha - D\nabla_\perp^2)(F_+ + F_-). \tag{141b}$$

From these results, we find that

$$F_{+,z} = -\left[\frac{1}{4\varepsilon D} + \varepsilon\alpha - D\nabla_\perp^2\right] F_+ + \left[\frac{1}{4\varepsilon D} - \varepsilon\alpha + D\nabla_\perp^2\right] F_-, \tag{142a}$$
$$F_{-,z} = \left[\frac{1}{4\varepsilon D} + \varepsilon\alpha - D\nabla_\perp^2\right] F_- - \left[\frac{1}{4\varepsilon D} - \varepsilon\alpha + D\nabla_\perp^2\right] F_+. \tag{142b}$$

 The system given in (142) gives the 3D two-flux equations. Perhaps this 3D extension to two-flux theory is useful for some applications, but since it is entirely consistent with the diffusion approximation, one may as well consider solving the 3D diffusion approximation instead.

10   Conclusions

We have discussed approximations to the RTE in strongly scattering medium. In particular, we have discussed boundary layer theory and two-flux theory. Both of these theories have connections to the DP_N method which we also explain. By investigating the connections between these two approximations and the DP_N method, we have gained valuable insight into each of them.

   Boundary layer theory gives a systematic perturbation method to compute the solution of the RTE in the limit as the albedo approaches one. It gives the solution as the sum of the interior solution and the boundary layer solutions. The interior solution is governed by the diffusion approximation. The boundary layer solutions satisfy a conservative, one-dimensional RTE in a half space. By requiring that boundary layer solutions satisfy asymptotic matching conditions, we derive boundary conditions for the diffusion approximation. Since we cannot solve the boundary layer problems analytically, we have used the DP_N method to solve them. This method also provides a convenient method to impose asymptotic matching which, in turn, provides a simple method for computing the coefficients in the boundary conditions for the diffusion approximation. The result of boundary layer theory is a uniformly valid asymptotic approximation that is accurate over the entire domain with a precisely defined error bound. Numerical results show that this asymptotic solution is very accurate over the entire domain.

Two-flux theory provides a simple and intuitive picture of light scattering and absorption in one spatial dimension. However, it has lacked a systematic theoretical framework to derive its coefficients and boundary conditions. Here, we have derived two-flux theory using the interior solution from boundary layer theory. By doing so, we have determined that two-flux theory is just a reformulation of the diffusion approximation. Consequently, it is accurate only for describing light that has penetrated deep into a strongly scattering medium. It is not accurate near boundaries or sources. This connection allows for the explicit determination of the coefficients. Specifically, we have found that the scattering coefficient is $S = 3\kappa_s(1 - g)/4 - \kappa_a(1 - 3g)/4$, and the absorption coefficient is $K = 2\kappa_a$. These results provide a useful connection between two-flux theory and the fundamental quantities in radiative transfer. However, just as the boundary conditions for the diffusion approximation can be problematic, so are those for two-flux theory. Here, we have derived boundary conditions for two-flux theory that are consistent with the results in boundary layer theory for the interior solution. By doing so, we have introduced the asymptotic two-flux theory. In contrast to naïve two-flux theory that uses the fluxes incident on the boundary as boundary data, the asymptotic two-flux theory has been shown to be more accurate in the interior of the domain away from boundaries. However, despite the fact that the naïve two-flux theory makes large errors in the interior of the domain, it approximates the fluxes exiting the domain at the boundaries better than the asymptotic two-flux theory.

Two-flux theory can be improved by adding another length scale that takes into account the rapid change of the solution near the boundaries. The result is called four-flux theory. Here, we have drawn a connection between four-flux theory and the DP_N method with $N = 1$. Numerical results show that four-flux theory computed from DP_1 provides an significant improvement over two-flux theory. However, four-flux theory is substantially less accurate than boundary layer theory.

Finally, we consider two-flux theory for three-dimensional radiative transfer problems. We have derived this three-dimensional two-flux theory from three-dimensional boundary layer theory just as we have done for one-dimensional problems. This result provides some insight into how two-flux theory extends to three-dimensional problems. However, since it is a reformulation of the diffusion approximation, it is perhaps unnecessary to consider it for practical problems.

The close connections discussed here between boundary layer theory, two-flux theory, and the DP_N method in strongly scattering media provides valuable insight into radiative transfer problems. Specifically, by drawing the connection between the diffusion approximation and two-flux theory, we have explicitly determined the coefficients in the two-flux equations. Additionally, the challenge in applying the diffusion approximation near boundaries should serve as caution in applying two-flux theory there, as well. Having these connections established should be useful to other researchers considering these approximations for problems of practical interest.

References

Aronson R (1986) P_N versus double-P_N approximations for highly anisotropic scattering. Transport Theor Stat 15(6–7):829–840

Bell GI, Glasstone S (1970) Nuclear reactor theory. Van Nostrand Reinhold Co, New York

Bender CM, Orszag SA (2013) Advanced mathematical methods for scientists and engineers I: asymptotic methods and perturbation theory. Springer Science & Business Media, Berlin

Brinkworth B (1972) Interpretation of the Kubelka-Munk coefficients in reflection theory. Appl Opt 11(6):1434–1435

Case KM, Zweifel PF (1967) Linear transport theory. Addison-Wesley, Reading

Chandrasekhar S (1960) Radiative transfer. Dover, New York

Courant R, Hilbert D (2008) Methods of mathematical physics, Vol. 2: differential equations. Wiley, New York

Dark JP, Kim AD (2017) Asymptotic theory of circular polarization memory. J Opt Soc Am A 34(9):1642–1650

Edström P (2007) Examination of the revised Kubleka-Munk theory: considerations of modeling strategies. J Opt Soc Am A 24(2):548–556

Gao H, Zhao H (2009) A fast-forward solver of radiative transfer equation. Transp Theory Stat Phys 38(3):149–192

Gate L (1974) Comparison of the photon diffusion model and Kubleka-Munk equation with the exact solution of the radiative transport equation. Appl Opt 13(2):236–238

Habetler G, Matkowsky B (1975) Uniform asymptotic expansions in transport theory with small mean free paths, and the diffusion approximation. J Math Phys 16(4):846–854

Hinch EJ (1991) Perturbation methods. Cambridge University Press, Cambridge

Ishimaru A (1999) Wave propagation and scattering in random media. Wiley-IEEE-Press, New York

Jensen HW (2001) Realistic image synthesis using photon mapping. A K Peters Ltd., Natick

Keller HB (1960) On the pointwise convergence of the discrete-ordinate method. J Soc Ind Appl Math 8(4):560–567

Kim AD (2011) Correcting the diffusion approximation at the boundary. J Opt Soc Am A 28(6):1007–1015

Kim AD, Keller JB (2003) Light propagation in biological tissue. J Opt Soc Am A 20(1):92–98

Kim AD, Moscoso M (2011) Diffusion of polarized light. Multiscale Model Simul 9(4):1624–1645

Kirk JT (1994) Light and photosynthesis in aquatic ecosystems. Cambridge University Press, Cambridge

Kokhanovsky A (2006a) Asymptotic radiative transfer. In: Light scattering reviews. Springer, Berlin, pp 253–289

Kokhanovsky AA (2006b) Cloud optics. Springer, Berlin

Kokhanovsky AA (2007) Physical interpretation and accuracy of the Kubelka-Munk theory. J Phys D 40(7):2210

Kubelka P (1948) New contributions to the optics of intensely light-scattering materials. Part I. J Opt Soc Am 38(5):448–457

Kubelka P, Munk F (1931) Ein Beitrag zur Optik der Farbanstriche. Z Tech Phys 12:593–601

Larsen EW (1999) The linear Boltzmann equation in optically thick systems with forward-peaked scattering. Prog Nucl Energy 34(4):413–423

Larsen EW, Keller JB (1974) Asymptotic solution of neutron transport problems for small mean free paths. J Math Phys 15(1):75–81

Lehtikangas O, Tarvainen T, Kim AD (2012) Modeling boundary measurements of scattered light using the corrected diffusion approximation. Biomed Opt Express 3(3):552–571

Lewis EE, Miller WF (1984) Computational methods of neutron transport. Wiley, New York

Malvagi F, Pomraning G (1991) Initial and boundary conditions for diffusive linear transport problems. J Math Phys 32(3):805–820

Marshak A, Davis A (2005) 3D radiative transfer in cloudy atmospheres. Springer Science & Business Media, Berlin

Miller PD (2006) Applied asymptotic analysis. American Mathematical Society, Providence

Mishchenko MI (2014) Electromagnetic scattering by particles and particle groups: an introduction. Cambridge University Press, Cambridge

Mishchenko MI, Travis LD, Lacis AA (2006) Multiple scattering of light by particles: radiative transfer and coherent backscattering. Cambridge University Press, Cambridge

Mobley CD (1994) Light and water: radiative transfer in natural waters. Academic Press, New York

Modest MF (2013) Radiative heat transfer. Academic Press, New York

Molenaar R, Jaap J, Zijp JR (1999) Determination of Kubelka-Munk scattering and absorption coefficients by diffuse illumination. Appl Opt 38(10):2068–2077

Myrick ML, Simcock MN, Baranowski M, Brooke H, Morgan SL, McCutcheon JN (2011) The Kubelka-Munk diffuse reflectance formula revisited. Appl Spectrosc Rev 46(2):140–165

Neuman M, Edström P (2010) Anisotropic reflectance from turbid media. I. theory. J Opt Soc Am A 27(5):1032–1039

Nobbs JH (1985) Kubelka-Munk theory and the prediction of reflectance. Color Technol 15(1):66–75

Peraiah A (2002) An introduction to radiative transfer: methods and applications in astrophysics. Cambridge University Press, Cambridge

Philips-Invernizzi B, Dupont D, Caze C (2001) Bibliographical review for reflectance of diffusing media. Opt Eng 40(6):1082–1092

Pomraning G, Ganapol B (1995) Asymptotically consistent reflection boundary conditions for diffusion theory. Ann Nucl Energy 22(12):787–817

Rohde SB, Kim AD (2012) Modeling the diffuse reflectance due to a narrow beam incident on a turbid medium. J Opt Soc Am A 29(3):231–238

Rohde S, Kim A (2014) Convolution model of the diffuse reflectance for layered tissues. Opt Lett 39(1):154–157

Rohde SB, Kim AD (2017) Backscattering of continuous and pulsed beams. Multiscale Model Simul 15(4):1356–1375

Şahin-Biryol D, Ilan B (2014) Asymptotic solution of light transport problems in optically thick luminescent media. J Math Phys 55(6):061, 501

Sandoval C, Kim AD (2014) Deriving Kubelka-Munk theory from radiative transport. J Opt Soc Am A 31(3):628–636

Schuster A (1905) Radiation through a foggy atmosphere. Astrophys J 21:1

Schwarzschild K (1906) On the equilibrium of the sun's atmosphere. Nach König Gesell Wiss Göttingen Math-Phys Klasse, 195, p 41–53 195:41–53

Sobolev VV (2017) Light scattering in planetary atmospheres: international series of monographs in natural philosophy, vol 76. Elsevier, Amsterdam

Star W, Marijnissen J, Van Gemert M (1988) Light dosimetry in optical phantoms and in tissues: I. multiple flux and transport theory. Phys Med Biol 33(4):437

Thennadil SN (2008) Relationship between the Kubleka-Munk scattering and radiative transfer coefficients. J Opt Soc Am A 25(7):1480–1485

Thomas GE, Stamnes K (2002) Radiative transfer in the atmosphere and ocean. Cambridge University Press, Cambridge

Tsang L, Kong JA, Shin RT (1985) Theory of microwave remote sensing. Wiley-Interscience, New York

Van de Hulst HC (2012) Multiple light scattering: tables, formulas, and applications. Elsevier, Amsterdam

Vargas WE, Niklasson GA (1997) Applicability conditions of the Kubelka-Munk theory. Appl Opt 36(22):5580–5586

Wang LV, Hi Wu (2012) Biomedical optics: principles and imaging. Wiley, New York

Welch AJ, Van Gemert MJ et al (2011) Optical-thermal response of laser-irradiated tissue, vol 2. Springer, Berlin

Yang L, Kruse B (2004) Revised Kubelka-Munk theory. I. Theory and application. J Opt Soc Am A 21(10):1933–1941

Yang L, Miklavcic SJ (2005) Revised Kubelka-Munk theory III. A general theory of light propagation in scattering and absorptive media. J Opt Soc Am A 22(9):1866–1873

Yang L, Kruse B, Miklavcic SJ (2004) Revised Kubelka-Munk theory. II. Unified framework for homogeneous and inhomogeneous optical media. J Opt Soc Am A 21(10):1942–1952

Polarized Radiation Transport Equation in Anisotropic Media

Check for updates

Margarita G. Kuzmina

The paper is dedicated to the memory of Evgraph Sergeevich Kuznetsov.

This paper is a translation of my old paper aimed at vector radiative transfer studies published back in 1976 in Russian as a preprint of Institute of Applied Mathematics of Russian Academy of Sciences (Kuzmina 1976). Because the paper has not been widely circulated, it remains almost unknown outside Russia. This is the main reason why I have been asked to prepare this translation for Springer Series in Light Scattering. Since 1976 excellent works have been published on this subject (Tsang and Ishimaru 1992; Tsang and Kong 1992; Mandt and Tsang 1992; Mishchenko et al. 2002; Mishchenko 2008a, b; Barabanenkov 2009; Mishchenko et al. 2016). However, the paper has its merits as it stands even now after more than 40 years after its publication. Therefore, it has been decided to prepare the slightly revised paper. Some misprints have been corrected. However, the general structure of the paper remains. Therefore, a reader can grasp a flavor of the approach to the problem existing back in the middle of last century.

In particular, in this paper polarized radiation transport equation in media composed of randomly spatially distributed discrete non-spherical scatterers, the scatterer sizes being comparable with the electromagnetic radiation wavelength, has been derived from the equations of classical electrodynamics. The transport equation derivation is based on the analysis of the system of equations for self-consistent radiation field that has been obtained as the solution to the problem of electromagnetic wave multiple scattering by an ensemble of isolated discrete scatterers. The electric field vector of coherently scattered electromagnetic wave (that propagates without scattering through continuous anisotropic medium) has been obtained via

M. G. Kuzmina (✉)
Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, Russia
e-mail: mg.kuzmina@gmail.com

statistical averaging of the governing system of equations. The optical properties of the anisotropic medium can be expressed in terms of the scattering amplitude operator of individual scatterers forming the dispersed medium and the scatterer spatial distribution. The non-coherently scattered radiation behavior is governed by vector radiation transport equation for anisotropic medium. The matrix extinction operator of the transport equation is expressed in terms of anisotropic medium optical properties whereas the scattering phase matrix is expressed through the law of scattering of individual scatterers. In addition to usual radiation transport processes the transport equation describes also some radiation interference phenomena related to different velocities of propagation and different absorptions of normal waves in anisotropic medium.

1 Introduction

In recent years a number of papers appeared where the vector transport equation describing polarized radiation transfer in optically isotropic and anisotropic media had been derived based on exact electrodynamics equations for multiple radiation scattering by statistical ensembles of scatterers.

In the frames of traditional approach to the problems of radiation multiple scattering the radiation transport equation is usually derived based on energy balance considerations. So the transport equation in fact represents linearized integro-differential Boltzmann equation. Since the Bolzmann equation exactly describes transport processes of massive particles in various media, the energy-based approach for transport equation obtaining implies that electromagnetic radiation is described in the frames of geometrical optics approximation. Therefore, as one should expect, in the frames of the traditional approach all the effects of radiation interference are neglected. And in addition, the approach of electromagnetic radiation description is admissible under the conditions $\lambda \ll l$, where l is the length of free radiation path between the acts of radiation scattering, λ being the radiation wavelength.

However, as was earlier known, for radiative transfer problems in some media (such as aerosol, fog and optically anisotropic media) the radiation interference phenomena cannot be completely neglected. So, a correct way of radiation transport equation deriving based on consideration of the problem of multiply scattered electromagnetic radiation by an ensemble of chaotically distributed discrete scatterers is desirable.

The problem was first considered and solved by Foldy (1945) for the case of dispersed medium composed of point uncorrelated isotropic-scatting scatterers, and the description of wave radiation field in the frames of scalar approximation. Although the paper (Foldy 1945) preceded the papers by Dyson (1949) and Salpeter and Bether (1951), it is just in the paper (Foldy 1945) the mathematically exact restrictions on radiation field and scattering medium were formulated, and the equations of Dyson- and Bether–Salpeter type were first obtained. The extension of the results by Foldy

to the case of dispersed medium composed of anisotropic-scattering correlated scatterers was further given by Lax (1951).

After attracting into transport theory of Dyson- and Bether–Salpeter equations, obtained via perturbation series summation, many papers devoted to obtaining and justification of scalar transport equation appeared. In the way the transport equations were obtained for random inhomogeneous media composed of discrete correlated scatterers (Gnedin and Dolginov 1964, 1965, 1967; Barabanenkov and Finkelberg 1967; Barabanenkov et al. 1969, 1970, 1971), for media composed of continuous scatterers formed by dielectric permittivity fluctuations (Gnedin et al. 1970a, b, c; Barabanenkov et al. 1970), and for media containing of both discrete and continuous scatterers (Ovchinnikov 1974). One of the ways of scalar transport equation deriving consisted in preliminary obtaining of so called generalized transport equation for spectral density of mean radiation field and the mean Green function, and consequent obtaining of usual transport equation from the generalized one (Ovchinnikov and Tatarsky 1972). Transport equation for spectral intensity tensor, accounting some interference effects and the phenomena caused by stimulated and spontaneous radiation, was derived in the frames of quantum-statistical approach (Kruglov 1978). The transport equation obtaining in the frames of quantum electrodynamics and consequent limit transfer to the classical radiation transport equation surely could be considered as the most exact approach. However, the approach based on classical radiation field description is in some sense preferable because of its simplicity and also because of comprehensive physical description of electromagnetic field properties. By the reason the approach developed in the paper (Watson 1969), where the solution to the problem of multiply scattered polarized radiation in rarefied electron plasma is given, seems quite adequate. In fact the approach presented in (Watson 1969) can be considered as the extension of approach developed in (Foldy 1945) to the case of vector radiation field and scatterer correlations accounting. Just as in the papers (Foldy 1945; Lax 1951), in the paper (Watson 1969) the coherently scattered radiation was selected, the medium refraction index was obtained, and radiation transport equation for coherence matrix was derived. The extension of the approach developed in (Watson 1969) to the case of turbulent plasma, where radiation propagates over curved rays, was further given in the paper (Law and Watson 1970). Yu. I. Gnedin, A. Z. Dolginov and N. A. Silant'ev derived polarized radiation transport equation for optically anisotropic media composed of correlated scatterers, small compared with radiation wavelength (Gnedin et al. 1970a, b; Dolginov et al. 1970; Silant'ev 1971). In particular, in the paper (Dolginov et al. 1970) the system of Lippman-Shwinger equations was first derived from the Maxwell equations, the intensity, the polarization and the radiation frequency change for multiply scattered radiation being further obtained.

In the presented paper the polarized radiation transport equation has been obtained for dispersed medium composed of discrete correlated scatterers, the scatterer sizes being comparable with radiation wavelength. The separation of scattered radiation field into coherently and non-coherently scattered components arises in the process of statistical averaging of multiply scattered radiation by the ensemble of discrete scatterers. The coherently scattered electromagnetic wave propagates without scat-

tering in effective continuous optically anisotropic medium. It is decomposed into the sum of two independent waves (normal waves) propagating in the medium with different phase velocities and different absorptions. The refraction indices, absorption coefficients and polarization states of the normal waves are expressed in terms of ensemble averaged operator of scattering amplitude in forward-scattering direction. The multiply scattered radiation transfer in the dispersed anisotropic medium is governed by vector transport equation. In the paper it is written for four-component polarization vector (in particular, for the Stokes vector). The scattering phase matrix of medium volume element is expressed in terms of scattering characteristics of individual scatterers and the scatterer statistical distribution. The differential operator of the vector transport equation reflects radiation interference effects caused by different phase velocities and different absorptions of normal waves in anisotropic medium.

2 Statement of the Problem. The System of Restrictions on Medium and Radiation Parameters

Let us consider the problem of multiple scattering of classical quasi-monochromatic electromagnetic radiation (of the wavelength λ) by sparse dispersed medium (an ensemble of N, $N \gg 1$, sparsely randomly distributed isolated macroscopic scatterers) occupying the volume V.

Suppose that

(01) Both spatial distribution of scatterers and the distribution on scattering characteristics (sizes, orientations etc.) are specified, the scattering characteristics being defined by a scattering parameter s (taking the values either from discrete or from continuous set of values).

(02) Scattering characteristics of individual scatterers are stationary and do not depend on the fact that a scatterer belongs to the dispersed medium.

(03) The velocities of scatterers are sufficiently small, so that spatial scatterer positions can be considered as adiabatically varying parameters (The condition is fulfilled if $|\mathbf{v}_{sc}| \ll v_{ph}$, where $\mathbf{v}_{ph}$ is the phase velocity of electromagnetic wave in the medium).

(04) The number of scatterers is great: $N \gg 1$.

Further the natural relations between the parameters of multiple scattering problem should be formulated in terms of the parameters—the length l of free radiation path between the acts of scattering, the average scatterer diameter d_s, the radiation wavelength λ In the case of very short wavelengths ($\lambda \ll l$) radiation transfer in the medium is governed by the Boltzmann equation. In the situation the radiation extinction is taken into account, whereas all the radiation interference effects are completely neglected. In the parametric domain $\lambda < l$ the geometrical optics approximation is valid. In the case the ray refraction in the process of multiple radiation scattering can be accounted, but the diffraction effects cannot be taken into account. So, the

radiation transfer problem can be formulated in terms of medium refractive index fluctuations (Gnedin et al. 1970a, b, c; Barabanenkov et al. 1972). At last, in the parametrical domain $\lambda \sim l$ both scatterer correlation effects become essential and the accounting of diffraction phenomena is of importance. Further problem simplification is possible under the condition $d_s \ll \lambda$ (when the approximation of elementary dipoles for scatterers is applicable, the case of Rayleigh scattering).

The approach to transport equation deriving used in the paper in fact represents the extension of the methods developed in the papers (Lax 1951; Watson 1969). It allows to consider the problem in the parametrical domain

$$\lambda < l, \quad d_s \geq \lambda, \quad d_s \ll l.$$

In addition we also admit that the condition

(05) $l \gg \lambda$ is satisfied as well (to avoid of bulky computations). The condition means that each new scattering act occurs in far-field zone of all the other scatterers.

We also admit the following additional simplifying restrictions:

(06) the coherent scattering occurs in exactly forward scattering direction (the restriction was not used in the paper (Waterman and Truell 1961)).

Indeed, as it is known (Watson 1969; Dolginov et al. 1970), the coherent scattering can be neglected outside the diffraction cone which angle is of the order $((\omega/c)R_s)^{-1}$, where R_s is the diameter of the whole volume of scattering medium. Therefore, the mentioned approximation is the more exactly satisfied, the better the inequality $(\omega/c)R_s \gg 1$ is fulfilled. The last inequality should be fulfilled sufficiently well for multiply scattering media, because otherwise (that is, at $(\omega/c)R_s \leq 1$ due to the condition (05) we would have $(\omega/c)l \gg 1$, and so the inequality $R_s \ll l$ would take place, that is, the approximation of single scattering would be valid in the medium).

(07) $R_c \ll l$, where R_c is the typical correlations length (the condition allows to be restricted by correlation functions of low orders).

The next three medium constraints are introduced for the sake of calculation simplification:

(08) the restriction $|n^{(\alpha)} - 1| \ll 1, \quad \alpha = 1, 2$, where $n^{(1,2)}$ are the eigenvalues of the medium refractive index operator (the consequence of medium sparsity).
(09) $\left| dn^{(\alpha)}/d\mathbf{r} \right| = 0, \quad \alpha = 1, 2$. (That is, the case of homogeneous medium is considered where the propagation of non-scattered radiation can be described in the frames of geometrical optics, the rays being straight lines).
(10) weak medium optical anisotropy approximation is accepted: $(\mathbf{E}(\mathbf{r}, \mathbf{k}), \mathbf{k}) = 0$, where $\mathbf{k}$ is the wave vector direction in the medium (that is, the transversality of electrical field vector of electromagnetic wave in the medium is admitted).

3   Multiple Scattering of Electromagnetic Wave by an Ensemble of Isolated Scatterers

Consider the problem of multiple scattering of plane quasi-monochromatic electromagnetic wave by an ensemble of N sparsely randomly distributed isolated macroscopic scatterers, the scatterer centers being defined by radius-vectors $\mathbf{r}_1, \ldots \mathbf{r}_N$, and scattering characteristics—by scattering parameters $s_1, \ldots s_N$. The electrical field vector of the electromagnetic wave in some spatial point $\mathbf{r}$ can be written as

$$\mathbf{E}^{inc}(\mathbf{r},\ \mathbf{k}_0,\ t) = \mathbf{E}^{inc}\ (\mathbf{r},\ \mathbf{k}_0)\, e^{-i\,\omega\, t} = \mathbf{E}^{inc}\, e^{i(\omega/c)\mathbf{k}_0\mathbf{r}}\, e^{-i\,\omega\, t}, \qquad (3.1)$$

where $k_0 = (\omega/c)\mathbf{k}_0$ is the wave vector of incident wave, $\mathbf{k}_0$ is the unit vector, normal to the wave front, $\mathbf{E}$ is the vector defining the electromagnetic wave polarization.

The electrical field of electromagnetic wave, scattered by the isolated scatterer with center location at the point $\mathbf{r}_n$, can be defined as the solution of the boundary-value problem for vector wave equation (Newton 1969), and in far-field zone of the scatterer can be written as

$$\mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}_n) = \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|)\hat{A}(\mathbf{r}_n,\ s_n;\ \mathbf{k}_0 \rightarrow \mathbf{k}_{n0})\, \mathbf{E}^{inc}(\mathbf{r}_n,\ \mathbf{k}_0), \qquad (3.2)$$

where

$$\mathbf{k}_{n0} = \frac{\mathbf{r} - \mathbf{r}_0}{|\mathbf{r} - \mathbf{r}_0|}, \quad k_0|\mathbf{r} - \mathbf{r}_0| \gg 1, \qquad (3.3)$$

$$\hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_0|) = (\hat{I} - \mathbf{k}_{n0}\mathbf{k}_{n0}) \cdot \frac{e^{i(\omega/c)|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r} - \mathbf{r}_0|}, \qquad (3.4)$$

and $\hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_0|)$ is the asymptotic expression in far-field zone for Green function of the vector wave equation. The linear operator $\hat{A}(\mathbf{r}_n, s_n; \mathbf{k}_0 \rightarrow \mathbf{k}_{n0})$ entering into the Eq. (3.2), is known as amplitude scattering operator. It defines the amplitude of divergent spherical scattered wave in far-field zone which can be written as

$$\mathbf{E}^{sc}(\mathbf{r}, \mathbf{k}_{n0}) = \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|)\mathbf{F}^{sc}(\mathbf{r}_n, \mathbf{k}_{n0}),$$
$$\mathbf{F}^{sc}(\mathbf{r}_n, \mathbf{k}_{n0}) \equiv_{df} \hat{A}(\mathbf{r}_n, s_n; \mathbf{k}_0 \rightarrow \mathbf{k}_{n0})\, \mathbf{E}^{inc}(\mathbf{r}_0, \mathbf{k}_0) \qquad (3.5)$$

In the case of system of N fixed scatterers the electrical field is the superposition of the waves scattered by all scatteres. Due to the condition (05) any scattered wave can be considered as locally plane one, and therefore all the wave vector directions are definitely specified. Indeed, let $\mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}_n)$ is the electric field vector at the point $\mathbf{r}$, scattered by the scatterer located at the point $\mathbf{r}_n$ and propagated further to the point $\mathbf{r}$ without scattering. Then one can write

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}_0) + \sum_{n=1}^{N} \mathbf{E}^{sc}(\mathbf{r}; \mathbf{r}_n), \qquad (3.6)$$

$$\mathbf{E}^{sc}(\mathbf{r}; \mathbf{r}_n) = \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|)\{\hat{A}(\mathbf{r}_n, s_n; \mathbf{k}_0 \rightarrow \mathbf{k}_{n0})\mathbf{E}^{inc}(\mathbf{r}_0, \mathbf{k}_0)$$

$$+ \sum_{m \neq n}^{N} \hat{A}(\mathbf{r}_n, s_n; \mathbf{k}_{mn} \rightarrow \mathbf{k}_{n0})\mathbf{E}^{sc}(\mathbf{r}_n; \mathbf{r}_m)\}, \ n = 1, \ldots N, \qquad (3.7)$$

where

$$\mathbf{k}_{n0} = \frac{\mathbf{r} - \mathbf{r}_n}{|\mathbf{r} - \mathbf{r}_n|}, \quad \mathbf{k}_{mn} \equiv_{df} \frac{\mathbf{r}_n - \mathbf{r}_m}{|\mathbf{r}_n - \mathbf{r}_m|}, \qquad (3.8)$$

Putting in Eq. (3.7) $\mathbf{r} = \mathbf{r}_k$ one can obtain the system of equations for $\mathbf{E}^{sc}(\mathbf{r}_k, \mathbf{r}_n)$ which can be considered as the basic system of equations for the problem of multiple scattering. The system (3.7) is intuitively clear because it accounts for each scatterer of the ensemble the fact of presence of all the other scatterers. The system was written first by Foldy (1945) for the ensemble of isotropic-scattering point scatterers. It was further extended by Lax (1951) to the case of ensemble of scatterers with arbitrary operator of scattering amplitude. Later Watson showed (Watson 1969) that the exact solution of vector wave equation satisfies the system (3.7) written for system of point dipoles.

The Eqs. (3.6) and (3.7) can be presented in more usual form after substitution the expression for $\mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}_n)$ into the right-hand part of the Eq. (3.7). The expansion of $\mathbf{E}(\mathbf{r})$ in the form of series on scattering multiplicity then can be seen:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}_0) + \sum_{n=1}^{N} \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|)\hat{A}(\mathbf{r}_n, s_n, \mathbf{k}_0 \rightarrow \mathbf{k}_{n0}) \mathbf{E}^{inc}(\mathbf{r}_n, \mathbf{k}_0)+$$

$$+ \sum_{n=1}^{N} \sum_{m \neq n}^{N} \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|)\hat{A}(\mathbf{r}_n, s_n, \mathbf{k}_{mn} \rightarrow \mathbf{k}_{n0})\hat{G}^{(\circ)}(|\mathbf{r}_n - \mathbf{r}_m|) \cdot$$

$$\cdot \hat{A}(\mathbf{r}_m, s_m, \mathbf{k}_0 \rightarrow \mathbf{k}_{mn})\mathbf{E}^{inc}(\mathbf{r}_m, \mathbf{k}_0) + \ldots \qquad (3.9)$$

The second term in Eq. (3.9) obviously corresponds to the contribution from single scattered incident wave by all the scatterers of the ensemble, the third term—to the contribution from secondly scattered wave and so on.

Define now the bases for calculation of electric field of electromagnetic wave that are necessary for polarization calculation. The matrix of amplitude scattering operator is usually defined in such a way that the matrix elements $A_{\alpha\beta}(\mathbf{r}, s; \mathbf{k}' \rightarrow \mathbf{k})$, $\alpha, \beta = 1, 2$ relate the components $E_\alpha^{inc}(\mathbf{r}, \mathbf{k}')$ of electrical field vector $\mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}')$ of electromagnetic wave, incident to the scatterer in the direction $\mathbf{k}'$ and calculated in some basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{k}'\}$, with the components $F_\beta^{sc}(\mathbf{r}, \mathbf{k})$, $\beta = 1, 2$ of the amplitude of electric field of the wave, scattered in the direction $\mathbf{k}$ and calculated in some basis

$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{k}\}$. We define these bases as $B^{(s')}(\mathbf{k}')$ and $B^{(s)}(\mathbf{k})$, correspondingly, and define their vectors by the formulas:

$$\begin{cases} \mathbf{e}_2^{(s')} = \mathbf{e}_2(\mathbf{k}') = \dfrac{\mathbf{k} \times \mathbf{k}'}{|\mathbf{k} \times \mathbf{k}'|}, \quad \mathbf{e}_1^{(s')} = \mathbf{e}_2^{(s')} \times \mathbf{k}', \\ \mathbf{e}_3^{(s')} = \mathbf{k}' = \mathbf{e}_1^{(s')} \times \mathbf{e}_2^{(s')}, \end{cases} \tag{3.10}$$

$$\begin{cases} \mathbf{e}_2^{(s)} = \mathbf{e}_2(\mathbf{k}) = \mathbf{e}_2^{(s')}, \quad \mathbf{e}_1^{(s)} = \mathbf{e}_2^{(s)} \times \mathbf{k}, \\ \mathbf{e}_3^{(s)} = \mathbf{k} = \mathbf{e}_1^{(s)} \times \mathbf{e}_2^{(s)}. \end{cases} \tag{3.11}$$

At $(\omega/c)\mathbf{k}(\mathbf{r} - \mathbf{r}_0) \gg 1$ the system of Eq. (3.7) can be written in components as follows:

$$\mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}_n) = \sum_{\alpha=1}^{2} E_\alpha^{(s)sc}(\mathbf{r}, \mathbf{k}_{n0})\mathbf{e}^{(s)}(\mathbf{k}_{n0}); \quad E_\alpha^{(s)sc}(\mathbf{r}, \mathbf{k}_{n0}) \equiv_{df} (\mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}_n), \mathbf{e}_\alpha^{(s)}(\mathbf{k}_{n0})),$$

$$\tag{3.12}$$

$$\mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}_0) = \sum_{\beta=1}^{2} E_\beta^{(s')inc}(\mathbf{r}, \mathbf{k}_0)\mathbf{e}^{(s')}(\mathbf{k}_0); \quad E_\beta^{(s')inc}(\mathbf{r}, \mathbf{k}_0) \equiv_{df} (\mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}_0), \mathbf{e}_\beta^{(s')}(\mathbf{k}_0)),$$

$$\tag{3.13}$$

$$E_\alpha^{(s)sc}(\mathbf{r}, \mathbf{k}_{n0}) = \frac{e^{i(\omega/c)|\mathbf{r}-\mathbf{r}_n|}}{|\mathbf{r} - \mathbf{r}_n|}\{\sum_{\beta=1}^{2} A_{\alpha\beta}(\mathbf{r}_n, s_n, \mathbf{k}_0 \to \mathbf{k}_{n0})E_\beta^{(s')inc}(\mathbf{r}_n, \mathbf{k}_0)$$

$$+ \sum_{m \neq n}^{N} \sum_{\beta'=1}^{2} A_{\alpha\beta'}(\mathbf{r}_n, s_n, \mathbf{k}_{mn} \to \mathbf{k}_{n0})E_{\beta'}^{(s')sc}(\mathbf{r}_n, \mathbf{k}_{mn})\}. \tag{3.14}$$

After introduction of the notations

$$\mathbf{E}^{(s)sc}(\mathbf{r}, \mathbf{k}_{n0}) \equiv_{df} \begin{bmatrix} E_1^{(s)sc}(\mathbf{r}, \mathbf{k}_{n0}) \\ E_2^{(s)sc}(\mathbf{r}, \mathbf{k}_{n0}) \end{bmatrix}$$

$$\hat{A}^{(s)}(\mathbf{r}, s, \mathbf{k}' \to \mathbf{k}) \equiv_{df} \begin{bmatrix} A_{11}(\mathbf{r}, s, \mathbf{k}' \to \mathbf{k}) & A_{12}(\mathbf{r}, s, \mathbf{k}' \to \mathbf{k}) \\ A_{21}(\mathbf{r}, s, \mathbf{k}' \to \mathbf{k}) & A_{22}(\mathbf{r}, s, \mathbf{k}' \to \mathbf{k}) \end{bmatrix},$$

the system of Eq. (3.14) can be rewritten in the form

$$\mathbf{E}^{(s)sc}(\mathbf{r}, \mathbf{k}_{n0}) = \frac{e^{i(\omega/c)|\mathbf{r}-\mathbf{r}_n|}}{|\mathbf{r} - \mathbf{r}_n|}\{\hat{A}^{(s)}(\mathbf{r}_n, s_n, \mathbf{k}_0 \to \mathbf{k}_{n0})\mathbf{E}^{(s')inc}(\mathbf{r}_n, \mathbf{k}_0)$$

$$+ \sum_{m \neq n}^{N} \hat{A}^{(s)}(\mathbf{r}_n, s_n, \mathbf{k}_{mn} \to \mathbf{k}_{n0})\mathbf{E}^{(s')sc}(\mathbf{r}_n, \mathbf{k}_{mn})\}. \tag{3.14*}$$

Besides the bases $B^{(s')}(\mathbf{k}')$ and $B^{(s)}(\mathbf{k})$ that are related to the plane of scattering it is necessary to introduce also the bases that would be naturally related to the dispersed medium. One of the proper bases is $B^{(0)}(\mathbf{k}) = \{\mathbf{e}^{(1)}(\mathbf{k}), \mathbf{e}^{(2)}(\mathbf{k})\}$ which vectors coincide with the polarization vectors of normal waves propagating in the direction $\mathbf{k}$ in the dispersed anisotropic medium. The vectors of the basis $B^{(0)}(\mathbf{k})$ satisfy the conditions

$$(\mathbf{e}^{(\alpha)}(\mathbf{k}), \mathbf{e}^{(\alpha)}(\mathbf{k})) = 1, \quad (\mathbf{e}^{(\alpha)}(\mathbf{k}), \mathbf{k}) = 0, \alpha = 1, 2; \quad (\mathbf{e}^{(1)}(\mathbf{k}), \mathbf{e}^{(2)}(\mathbf{k})) = 0. \quad (3.15)$$

Another basis of real-valued vectors that is convenient for Stokes parameters calculation, can be defined as follows. Let $\mathbf{n}_0$ be some physically selected direction in the medium. Define the vectors of the basis $B^{(l)}(\mathbf{k}) = \{\mathbf{e}_1^{(l)}(\mathbf{k}), \mathbf{e}_2^{(l)}(\mathbf{k}), \mathbf{e}_3^{(l)}(\mathbf{k})\}$ by the formulas

$$\begin{cases} \mathbf{e}_2^{(l)}(\mathbf{k}) = \dfrac{\mathbf{n}_0 \times \mathbf{k}}{|\mathbf{n}_0 \times \mathbf{k}|}, \quad \mathbf{e}_1^{(l)} = \mathbf{e}_2^{(l)} \times \mathbf{k}, \\ \mathbf{e}_3^{(l)} = \mathbf{k} = \mathbf{e}_1^{(l)} \times \mathbf{e}_2^{(l)}, \end{cases} \quad (3.16)$$

The components of vector $\mathbf{E}^{sc}(\mathbf{r}, \mathbf{k})$ in different bases are connected by the known linear transformations. For example, if we put

$$\mathbf{E}^{(l)sc}(\mathbf{r}, \mathbf{k}) = \hat{\mathcal{L}}(\mathbf{k})\mathbf{E}^{(s)sc}(\mathbf{r}, \mathbf{k}), \quad (3.17)$$

where

$$\mathcal{L}_{\alpha\beta}(\mathbf{k}) = (\mathbf{e}_\alpha^{(s)}(\mathbf{k}), \mathbf{e}_\beta^{(l)}(\mathbf{k})), \quad \alpha, \beta = 1, 2, \quad (3.18)$$

then the matrix of amplitude operator, defining the relation between the components of the vectors $\mathbf{F}^{(l)sc}(\mathbf{r}, \mathbf{k})$ and $\mathbf{F}^{(0)sc}(\mathbf{r}, \mathbf{k})$, is expressed through the matrix $\hat{A}^{(s)}(\mathbf{r}, s; \mathbf{k}' \to \mathbf{k})$ by the known formula

$$\hat{A}^{(l)}(\mathbf{r}, s, \mathbf{k}' \to \mathbf{k}) = \hat{\mathcal{L}}(\mathbf{k})\hat{A}^{(s)}(\mathbf{r}, s, \mathbf{k}' \to \mathbf{k})\hat{\mathcal{L}}^{-1}(\mathbf{k}') \quad (3.19)$$

4 Statistical Description of the Medium. Configuration Averaging

Following the method developed in (Foldy 1945; Lax 1951; Watson 1969), we define configuration ensemble of scatterers by introduction of the probability density $\mathcal{P}_N(\mathbf{r}_1, \ldots \mathbf{r}_N; s_1, \ldots s_N)$ so that $\mathcal{P}_N \Pi_{k=1}^{N} d\mathbf{r}_k ds_k$ is the probability to find N scatterers inside the volume element $\Pi_{k=1}^{N} d\mathbf{r}_k ds_k$ of the $3N \cdot N$-dimensional phase space. Obviously, the functions $\mathcal{P}_N(\mathbf{r}_1, \ldots \mathbf{r}_N; s_1, \ldots s_N)$ should satisfy the normalization condition

$$\int\limits_{V} \ldots \int\limits_{V} \int\limits_{s_{\min}}^{s_{\max}} \ldots \int\limits_{s_{\min}}^{s_{\max}} \mathcal{P}_N(\mathbf{r}_1, \ldots \mathbf{r}_N; s_1, \ldots s_N) \Pi_{k=1}^{N} d\mathbf{r}_k ds_k = 1. \qquad (4.1)$$

The probability density for single scatterer with given scattering properties is defined by

$$\mathcal{P}_1(\mathbf{r}_j, s_j) = \int\limits_{V} \ldots \int\limits_{V} \int\limits_{s_{\min}}^{s_{\max}} \ldots \int\limits_{s_{\min}}^{s_{\max}} \mathcal{P}_N(\mathbf{r}_1, \ldots \mathbf{r}_N; s_1, \ldots s_N) \Pi_{k \neq j}^{N} d\mathbf{r}_k ds_k, \qquad (4.2)$$

and for given scatterer pair—by the expression

$$\mathcal{P}_2(\mathbf{r}_j, s_j) = \int\limits_{V} \ldots \int\limits_{V} \int\limits_{s_{\min}}^{s_{\max}} \ldots \int\limits_{s_{\min}}^{s_{\max}} \mathcal{P}_N(\mathbf{r}_1, \ldots \mathbf{r}_N; s_1, \ldots s_N) \Pi_{k \neq i,j}^{N} d\mathbf{r}_k ds_k. \qquad (4.3)$$

For statistically independent scatterers we obviously have

$$\mathcal{P}_N(\mathbf{r}_1, \ldots \mathbf{r}_N; s_1, \ldots s_N) = \Pi_{k=1}^{N} \mathcal{P}_1(\mathbf{r}_j, s_j), \qquad (4.4)$$

and the difference

$$\mathcal{P}_2(\mathbf{r}_i, \mathbf{r}_j; s_i, s_j) - \mathcal{P}_1(\mathbf{r}_i, s_i)\mathcal{P}_1(\mathbf{r}_j, s_j)$$

characterizes the correlation degree between the scatterers. For ensemble of correlated scatterers it is convenient to express $\mathcal{P}_2(\mathbf{r}_i, \mathbf{r}_j; s_i, s_j)$ in terms of correlation functions $g_2(\mathbf{r}_i, \mathbf{r}_j; s_i, s_j)$:

$$\mathcal{P}_2(\mathbf{r}_i, \mathbf{r}_j; s_i, s_j) = \mathcal{P}_1(\mathbf{r}_i, s_i)\mathcal{P}_1(\mathbf{r}_j, s_j)\{1 + g_2(\mathbf{r}_i, \mathbf{r}_j; s_i, s_j)\}. \qquad (4.5)$$

We will use further the approximation

$$g_2(\mathbf{r}_i, \mathbf{r}_j; s_i, s_j) = 0 \text{ at } |\mathbf{r}_i - \mathbf{r}_j| > R_c. \qquad (4.6)$$

The number of scatterers of all types in unit medium volume, located in the vicinity of spatial point $\mathbf{r}$, is defined by

$$n(\mathbf{r}) = \int\limits_{s_{\min}}^{s_{\max}} \mathcal{P}_1(\mathbf{r}, s) ds. \qquad (4.7)$$

The density of scatterers with scattering parameter s, and their correlation density are given by the expressions

$$\rho(\mathbf{r}, s) = N\mathcal{P}_1(\mathbf{r}, s), \tag{4.8}$$

$$n(\mathbf{r}, \mathbf{r}'; s, s') = N^2\mathcal{P}_2(\mathbf{r}, \mathbf{r}'; s, s'). \tag{4.9}$$

It is convenient to introduce the conditional probabilities for calculation of configuration averages:

$$
\begin{aligned}
\mathcal{P}_N(\mathbf{r}_1, \dots \mathbf{r}_N; s_1, \dots s_N) &= \mathcal{P}_1(\mathbf{r}_i, s_i)\mathcal{P}^{(i)}_{N-1}(\mathbf{r}_1, \dots \mathbf{r}_N; s_1, \dots s_N / \mathbf{r}_i, s_i) \\
&= \mathcal{P}_2(\mathbf{r}_i, \mathbf{r}_j; s_i, s_j)\mathcal{P}^{(i,j)}_{N-2}(\mathbf{r}_1, \dots \mathbf{r}_N; s_1, \dots s_N / \mathbf{r}_i, \mathbf{r}_j; s_i, s_j).
\end{aligned}
\tag{4.10}
$$

Then configuration average for some random function $F(\mathbf{r}_1, \dots \mathbf{r}_N; s_1, \dots s_N)$ can be expressed as

$$
\begin{aligned}
\langle F(\mathbf{r}) \rangle \equiv_{df} \int\limits_V \dots \int\limits_V \int\limits_{s_{\min}}^{s_{\max}} \dots \int\limits_{s_{\min}}^{s_{\max}} F(\mathbf{r}_1, \dots \mathbf{r}_N; s_1, \dots s_N) \times \\
\times \mathcal{P}_N(\mathbf{r}_1, \dots \mathbf{r}_N; s_1, \dots s_N)\Pi^N_{k=1}d\mathbf{r}_k ds_k.
\end{aligned}
\tag{4.11}
$$

The conditional averages are defined by the usual formulas

$$\langle F(\mathbf{r}) \rangle_i \equiv_{df} \int\limits_V \dots \int\limits_V \int\limits_{s_{\min}}^{s_{\max}} \dots \int\limits_{s_{\min}}^{s_{\max}} F(\mathbf{r}_1, \dots \mathbf{r}_N; s_1, \dots s_N)\mathcal{P}^{(i)}_{N-1}\Pi^N_{k \neq i}d\mathbf{r}_k ds_k, \tag{4.12}$$

and the relation

$$\langle F(\mathbf{r}) \rangle = \int\limits_V \int\limits_{s_{\min}}^{s_{\max}} \langle F(\mathbf{r}) \rangle_i \mathcal{P}_1(\mathbf{r}_i, s_i)d\mathbf{r}_i ds_i \tag{4.13}$$

obviously takes place. If the correlation between the scatterer locations and its states is absent, one can write

$$\rho(\mathbf{r}, s) = \rho(\mathbf{r})q(s),$$

where $\rho(\mathbf{r})$ is the average scatterer density (independently on the scatterer sort).

5 Coherently Scattered Radiation. The Relation Between Medium Optical Parameters and Scattering Characteristics of Individual Scatterers

The system of Eqs. (3.6)–(3.7) allows, in principle, to exactly solve the problem of electromagnetic wave multiple scattering by arbitrary ensemble of scatterers. However, since N is very great, only statistically averaged solution (over scatterer con-

figurations and scattering characteristics) is of interest. We perform the averaging of the system (3.6) in the frames of accepted restrictions. The approximate integral equation for averaged electromagnetic field in the medium (the field of coherently scattered wave) can be obtained as the result.

By performing the averaging procedure of both the sides of the Eq. (3.7) (with accounting (4.12)–(4.13)) one can find

$$\langle \mathbf{E}^{sc}(\mathbf{r}) \rangle = \sum_{n}^{N} \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|)\{\langle \hat{A}(\mathbf{r}_n, s_n; \mathbf{k}_0 \to \mathbf{k}_{n0}) \rangle_n \mathbf{E}^{inc}(\mathbf{r}_n, \mathbf{k}_0)$$

$$+ \sum_{n}^{N} \langle \hat{A}(\mathbf{r}_n, s_n; \mathbf{k}_{mn} \to \mathbf{k}_{n0}) \mathbf{E}^{sc}(\mathbf{r}_n; \mathbf{r}_m) \rangle_{mn} \}. \tag{5.1}$$

Under the accepted approximation that coherent scattering occurs only in exact forward direction, (see the condition (06)) in the simplest case of the ensemble of statistically independent scatterers the Eq. (5.1) can be reduced to the equation

$$\langle \mathbf{E}^{sc}(\mathbf{r}, \mathbf{k}_0) \rangle = \int\limits_{V} \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}'|)\hat{\mathcal{A}}(\mathbf{r}'; \mathbf{k}_0 \to \mathbf{k}_0)\{\mathbf{E}^{inc}(\mathbf{r}', \mathbf{k}_0) + \sum_{n}^{N} \langle \mathbf{E}^{sc}(\mathbf{r}', \mathbf{k}_0) \rangle_{\mathbf{r}'} \}d\mathbf{r}',$$
$$\tag{5.2}$$

where $\langle \mathbf{E}^{sc}(\mathbf{r}', \mathbf{k}_0) \rangle_{\mathbf{r}'}$ is the field of scattered wave in the vicinity of scatterer located at the point $\mathbf{r}'$, averaged over the locations of all the scatterers besides the scatterer itself.

Following (Watson 1969), the additional approximation in the process of averaging of Eq. (3.7) was also used:

$$\langle \mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}') \rangle_{\mathbf{r}'} \approx \langle \mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}') \rangle. \tag{5.3}$$

From physical viewpoint it means that the radiation field is only slightly different from the field that would occur at the vicinity of the point $\mathbf{r}'$ if the scatterer be absent at the point. Obviously, the approximation (5.3) is fulfilled well for very large scatterer ensembles (at $N \gg 1$). For ensemble of correlated scatterers the simplest approximation can be written as (Lax 1951)

$$\langle \mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}') \rangle_{\mathbf{r}'} \approx c \langle \mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}') \rangle, \tag{5.4}$$

where the constant c is defined by the mean level of correlations. Accepting the approximation (5.3) and setting

$$\mathbf{E}^c(\mathbf{r}, \mathbf{k}_0) = \mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}_0) + \langle \mathbf{E}^{sc}(\mathbf{r}, \mathbf{k}_0) \rangle, \tag{5.5}$$

one can obtain the following integral equation for $\mathbf{E}^c(\mathbf{r}, \mathbf{k}_0)$:

$$\mathbf{E}^c(\mathbf{r}, \mathbf{k}_0) = \mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}_0) + \int d\mathbf{r}' \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}'|)\hat{\mathcal{A}}(\mathbf{r}'; \mathbf{k}_0 \to \mathbf{k}_0)\mathbf{E}^c(\mathbf{r}', \mathbf{k}_0), \quad (5.6)$$

where

$$\hat{\mathcal{A}}(\mathbf{r}; \mathbf{k}_0 \to \mathbf{k}_0) = c \int\limits_{s_{\min}}^{s_{\max}} \rho(\mathbf{r}, s)\hat{A}(\mathbf{r}, s; \mathbf{k}_0 \to \mathbf{k}_0)ds. \quad (5.7)$$

Applying the operator $\nabla \times \nabla \times \mathbf{f} + (\omega/c)^2 \mathbf{f}$ to both the sides of the Eq. (5.6) and using the equation for Green function $\hat{G}(|\mathbf{r} - \mathbf{r}'|)$, one will found that $\mathbf{E}^c(\mathbf{r}, \mathbf{k}_0)$ satisfies the differential equation

$$\nabla \times \nabla \times \mathbf{E}^c(\mathbf{r}, \mathbf{k}_0) + (\omega/c)^2 \hat{n}^2(\mathbf{r}, \mathbf{k}_0)\mathbf{E}^c(\mathbf{r}, \mathbf{k}_0) = 0, \quad (5.8)$$

where

$$\hat{n}^2(\mathbf{r}, \mathbf{k}_0) \equiv_{df} \hat{I} + 4\pi(\omega/c)^{-2}\hat{\mathcal{A}}(\mathbf{r}; \mathbf{k}_0 \to \mathbf{k}_0). \quad (5.9)$$

So, the coherently scattered radiation field satisfy the wave equation for continuous optically anisotropic medium with the refractive index operator $\hat{n}(\mathbf{r}, \mathbf{k}_0)$ that is defined by statistically averaged operator of scattering amplitude of individual scatterers (concerning the properties of $\hat{n}(\mathbf{r}, \mathbf{k}_0)$ see (Newton 1969), for example). For sparse dispersed medium, when

$$|n_{\alpha\beta}(\mathbf{r}, \mathbf{k}_0) - 1| \ll 1, \quad \alpha, \beta = 1, 2, \quad (5.10)$$

one can obtain from Eq. (5.9):

$$\hat{n}(\mathbf{r}, \mathbf{k}_0) \approx \hat{I} + 2\pi(\omega/c)^{-2}\hat{\mathcal{A}}(\mathbf{r}; \mathbf{k}_0 \to \mathbf{k}_0). \quad (5.11)$$

If scatterer spatial distribution and the distribution on scattering characteristics are not correlated, that is, when

$$\rho(\mathbf{r}, s) = \rho(\mathbf{r})q(s),$$

the operator $\hat{n}(\mathbf{r}, \mathbf{k}_0)$ can be expressed by the formula

$$\hat{n}(\mathbf{r}, \mathbf{k}_0) \approx \hat{I} + 2\pi(\omega/c)^{-2}\rho(\mathbf{r}) \int\limits_{s_{\min}}^{s_{\max}} \hat{A}(\mathbf{r}, s; \mathbf{k}_0 \to \mathbf{k}_0)ds. \quad (5.11a)$$

So, for homogeneous medium we have

$$\hat{n}(\mathbf{k}_0) \approx \hat{I} + 2\pi(\omega/c)^{-2}\rho \int_{s_{\min}}^{s_{\max}} \hat{A}(s;\mathbf{k}_0 \to \mathbf{k}_0)ds. \qquad (5.11b)$$

At last, if the scatterer ensemble consists of discrete collection of M scatterer sorts each being distributed with constant density, then

$$\hat{n}(\mathbf{k}_0) \approx \hat{I} + 2\pi(\omega/c)^{-2}\sum_{m=1}^{M} N_m \hat{A}_m(\mathbf{k}_0 \to \mathbf{k}_0). \qquad (5.11c)$$

where N_m is the number of scatterers of sort m in the unit volume. Therefore, in general case the coherently scattered electromagnetic wave consists of two independent waves:

$$\mathbf{E}^c(\mathbf{r},\mathbf{k}_0) = \sum_{\alpha=1}^{2} E^{(\alpha)}\mathbf{e}^{(\alpha)}(\mathbf{k}_0)e^{i(\omega/c)(n^{(\alpha)}+i\kappa^{(\alpha)})}, \quad E^{(\alpha)} = (E^{inc},\mathbf{e}^{(\alpha)}), \quad \alpha = 1,2.$$
$$(5.12)$$

The polarization vectors of normal waves $\mathbf{e}^{(1,2)}(\mathbf{k}_0)$ (not orthogonal in general case), their refractive indices $n^{(1,2)}(\mathbf{k}_0)$, and absorption coefficients $\kappa^{(1,2)}(\mathbf{k}_0)$ are defined by the eigenvalues $\hat{\mathcal{A}}^{(1,2)}$ and the eigenvectors $\mathbf{e}^{(1,2)}(\mathbf{k}_0)$ of the operator $\hat{\mathcal{A}}(\mathbf{r},\mathbf{k}_0 \to \mathbf{k}_0)$(see Supplement):

$$\mathrm{n}^{(\alpha)}(\mathbf{k}_0) = 1 + 2\pi(\omega/c)^{-2}\mathrm{Re}\mathcal{A}^{(\alpha)}(\mathbf{r};\mathbf{k}_0 \to \mathbf{k}_0), \qquad (5.13)$$

$$\kappa^{(\alpha)}(\mathbf{k}_0) = 2\pi(\omega/c)^{-2}\mathrm{Im}\mathcal{A}^{(\alpha)}(\mathbf{r};\mathbf{k}_0 \to \mathbf{k}_0), \quad \alpha = 1,2. \qquad (5.14)$$

The formulas (5.11) allow to analyze the dependence of medium optical characteristics on the features of scatterer ensemble. Consider, for example, a slab of dispersed medium composed of identical scatterers oriented at the same angle respectively the slab normal and completely chaotically distributed over the azimuth angle. Then simple qualitative considerations give the result: linear birefringence and dichroism cannot be inherent to the dispersed medium. On the contrary, the medium can possess only circular birefringence and dichroism in the case. In general, optical anisotropy of dispersed medium composed of completely chaotically distributed non-spherical scatterers can arise only in the case of scatterers of special shape (with broken mirror symmetry).

Accounting of correlations leads to appearance of additional term in the formula (5.9). For dispersed medium composed of point dipoles the additional correction term in the formula for medium refractive index (arising due to pair scatterer correlations) was obtained by Watson (1969). The method of calculating of the correction term to $\hat{n}^2(\mathbf{k})$ due to pair correlations, developed in (Watson 1969), can be extended to the case of scatterers with arbitrary amplitude scattering operator (it is performed at present paper, see Eqs. (7.14), (7.18)).

6 System of Equations for Non-coherently Scattered Radiation

To calculate the non-coherently multiply scattered radiation field defined by Eq. (5.12) it is helpful to pass from the system of Eqs. (3.6)–(3.7) to the new one, where the decomposition of multiply scattered electromagnetic wave into the sum of coherently and non-coherently scattered waves would be performed. Let $\mathbf{F}^{sc}(\mathbf{r}, \mathbf{r}_n)$ be the electrical field vector of non-coherently scattered wave by the scatterer located at the point $\mathbf{r}_n$ and transferred to the point $\mathbf{r}$ after undergoing coherent scattering only. Then one can write:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^c(\mathbf{r}, \mathbf{k}_0) + \sum_{n=1}^{N} \mathbf{F}^{sc}(\mathbf{r}, \mathbf{r}_n) = \mathbf{E}^c(\mathbf{r}, \mathbf{k}_0) + \mathbf{F}^{sc}(\mathbf{r}), \tag{6.1}$$

$$\mathbf{F}^{sc}(\mathbf{r}) \equiv_{df} \sum_{n=1}^{N} \mathbf{E}^{sc}(\mathbf{r}, \mathbf{r}_n) - \{\mathbf{E}^{sc}(\mathbf{r}, \mathbf{k}_0) - \mathbf{E}^{inc}(\mathbf{r}, \mathbf{k}_0)\},$$

$$\langle \mathbf{F}^{sc}(\mathbf{r}) \rangle = 0. \tag{6.2}$$

The functions $\mathbf{F}^{sc}(\mathbf{r}, \mathbf{r}_n)$ satisfy the system of equations

$$\mathbf{F}^{sc}(\mathbf{r}, \mathbf{r}_n) = \hat{G}(|\mathbf{r} - \mathbf{r}_n|; \mathbf{k}_{n0})\{\hat{\mathbf{A}}(\mathbf{r}_n, s_n; \mathbf{k}_0 \rightarrow \mathbf{k}_{n0})\mathbf{E}^c(\mathbf{r}_n, \mathbf{k}_0)$$

$$+ \sum_{n}^{N} \hat{\mathbf{A}}(\mathbf{r}_n, s_n; \mathbf{k}_{n0} \rightarrow \mathbf{k}_{mn})\mathbf{F}^{sc}(\mathbf{r}_n, \mathbf{r}_m)\}, \tag{6.3}$$

where $\hat{G}(|\mathbf{r}|, \mathbf{k})$ is the asymptotic expression for Green function of the wave Eq. (5.8) in infinite continuous anisotropic medium in far field zone (Newton 1969; Law and Watson 1970). In the case of homogeneous anisotropic medium the elements of matrix-function $\hat{G}(|\mathbf{r}|, \mathbf{k})$ are defined by simple expressions. Being calculated in the basis $\{\mathbf{e}^{(1)}(\mathbf{k}), \mathbf{e}^{(2)}(\mathbf{k}), \mathbf{k}\}$, where $\mathbf{e}^{(\alpha)}(\mathbf{k})$, $\alpha = 1, 2$ are the eigenvectors of the operator $\hat{\mathcal{A}}(\mathbf{r}, \mathbf{k}_0 \rightarrow \mathbf{k}_0)$, and $n^{(\alpha)}(\mathbf{k})$, $\alpha = 1, 2$ are defined by its eigenvalues accordingly to (5.13), the matrix $\hat{G}(|\mathbf{r}|, \mathbf{k})$ is the diagonal matrix with the elements

$$\hat{G}_{\alpha\alpha}(|\mathbf{r}|, \mathbf{k}) = r^{-1}e^{i(\omega/c)n^{(\alpha)}(\mathbf{k})r}, \quad \alpha = 1, 2, \quad \hat{G}_{33}(|\mathbf{r}|, \mathbf{k}) = 0. \tag{6.4}$$

Following (Watson 1969), the system of Eq. (6.3) can be derived from (5.2) via using the series on multiply scattered waves and performing the configuration averaging procedure. As a result we obtain the following decomposition for $\mathbf{F}^{sc}(\mathbf{r}; \mathbf{r}_n)$:

$$\mathbf{F}^{sc}(\mathbf{r}; \mathbf{r}_n) = \sum_{n}^{N} \{ \sum_{\alpha=1}^{2} G^{(\alpha)}(|\mathbf{r} - \mathbf{r}_n|; \mathbf{k}_{n0}) \mathbf{e}^{(\alpha)}(\mathbf{k}_{n0}) \sum_{\alpha=1}^{2} A^{(\circ)\alpha\beta}(\mathbf{r}_n, s_n; \mathbf{k}_0 \to \mathbf{k}_{n0}) E^{c\beta}(\mathbf{r}_0, \mathbf{k}_0)$$

$$+ \sum_{\alpha=1}^{2} G^{(\alpha)}(|\mathbf{r} - \mathbf{r}_n|; \mathbf{k}_{n0}) \mathbf{e}^{(\alpha)}(\mathbf{k}_{n0}) \times$$

$$\times \sum_{m \neq n}^{N} \sum_{\beta'=1}^{2} A^{(\circ)\alpha\beta'}(\mathbf{r}_n, s_n; \mathbf{k}_{mn} \to \mathbf{k}_{n0}) F^{(\circ)\beta'}(\mathbf{r}_n, \mathbf{r}_m; \mathbf{k}_{mn}) \}$$

$$= \sum_{n=1}^{N} \sum_{\alpha=1}^{2} F^{(\circ)\alpha}(\mathbf{r}, \mathbf{r}_n; \mathbf{k}_{n0}) \mathbf{e}^{(\alpha)}(\mathbf{k}_{n0}). \tag{6.5}$$

Defining $\mathbf{F}^{(\circ)}(\mathbf{r}, \mathbf{r}_n; \mathbf{k}_{n0})$ as

$$\mathbf{F}^{(\circ)}(\mathbf{r}, \mathbf{r}_n; \mathbf{k}_{n0}) \equiv_{df} \begin{bmatrix} (\mathbf{F}^{sc}(\mathbf{r}, \mathbf{r}_n), \mathbf{e}^{(1)}(\mathbf{k}_{n0})) \\ (\mathbf{F}^{sc}(\mathbf{r}, \mathbf{r}_n), \mathbf{e}^{(2)}(\mathbf{k}_{n0})) \end{bmatrix}, \tag{6.6}$$

we have

$$\mathbf{F}^{(\circ)}(\mathbf{r}, \mathbf{r}_n; \mathbf{k}_{n0}) = \hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|, \mathbf{k}_{n0}) \{ \hat{A}^{(\circ)}(\mathbf{r}_n, s_n; \mathbf{k}_0 \to \mathbf{k}_{n0}) \mathbf{E}^{(\circ)c}(\mathbf{r}_n, \mathbf{k}_0)$$

$$+ \sum_{m \neq n}^{N} \hat{A}^{(\circ)}(\mathbf{r}_n, s_n; \mathbf{k}_{mn} \to \mathbf{k}_{n0}) \mathbf{F}^{(\circ)}(\mathbf{r}_n, \mathbf{r}_m; \mathbf{k}_{mn}) \}, \tag{6.7}$$

where the matrices $\hat{A}^{(\circ)}(\mathbf{r}_n, s_n; \mathbf{k}_0 \to \mathbf{k}_{n0})$, $\hat{A}^{(\circ)}(\mathbf{r}_n, s_n; \mathbf{k}_{mn} \to \mathbf{k}_{n0})$, and the columns $\mathbf{E}^c(\mathbf{r}, \mathbf{k}_0)$, $\mathbf{F}^{(\circ)}(\mathbf{r}_n, \mathbf{r}_m; \mathbf{k}_{mn})$ are defined by formulas (3.17), (3.19), and $\hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}_n|, \mathbf{k}_{n0})$ is the matrix of second order with the elements

$$G^{(\circ)\alpha\beta}(|\mathbf{r} - \mathbf{r}_n|, \mathbf{k}_{n0}) = G^{(\alpha)}(|\mathbf{r} - \mathbf{r}_n|, \mathbf{k}_{n0}) \delta_{\alpha\beta}, \quad G^{(\alpha)} = |\mathbf{r} - \mathbf{r}_n|^{-1} e^{i(\omega/c) n^{(\alpha)}(\mathbf{k}_{n0})|\mathbf{r} - \mathbf{r}_n|}. \tag{6.8}$$

7 Integral Radiation Transport Equation

Radiation transport equation for scattering medium is usually written either for two-dimensional coherence matrix $\hat{J}(\mathbf{r}, \mathbf{k})$ or for four-component polarization vector $\mathbf{I}(\mathbf{r}, \mathbf{k}) = [I_1(\mathbf{r}, \mathbf{k}), I_2(\mathbf{r}, \mathbf{k}), I_3(\mathbf{r}, \mathbf{k}), I_4(\mathbf{r}, \mathbf{k})]^T$ which components are linear combinations of the matrix $\hat{J}(\mathbf{r}, \mathbf{k})$ (for example, for Stokes vector $\mathbf{I}^S(\mathbf{r}, \mathbf{k}) = [I(\mathbf{r}, \mathbf{k}), Q(\mathbf{r}, \mathbf{k}), U(\mathbf{r}, \mathbf{k}), V(\mathbf{r}, \mathbf{k})]^T$). The elements of the coherence matrix $\hat{J}(\mathbf{r}, \mathbf{k})$ are calculated in some basis $\{\mathbf{e}_1^{(l)}(\mathbf{k}), \mathbf{e}_2^{(l)}(\mathbf{k})\}$, that is naturally related to the medium and satisfies the conditions $(\mathbf{e}_\alpha^{(l)}, \mathbf{k}) = 0, \alpha = 1, 2$. The elements $J_{\alpha\alpha}(\mathbf{r}, \mathbf{k})$ of the $\hat{J}(\mathbf{r}, \mathbf{k})$ define the values of radiation flux density propagating along the direction $\mathbf{k}$ inside unit solid angle and possessing the polarization states $\mathbf{e}_\alpha^{(l)}(\mathbf{k})$. If

$(\mathbf{e}_\alpha^{(l)}, \mathbf{e}_\beta^{(l)}) = \delta_{\alpha\beta}$, the value $I(\mathbf{r}, \mathbf{k}) = \mathbf{Tr} \quad \hat{J}(\mathbf{r}, \mathbf{k})$ is the radiation intensity, propagating along the direction $\mathbf{k}$ inside unit solid angle. Then we have

$$J_{\alpha\beta}(\mathbf{r}, \mathbf{k}) = (c/8\pi)\langle E_\alpha^{(l)}(\mathbf{r}, \mathbf{k})E_\beta^{(l)}(\mathbf{r}, \mathbf{k})\rangle \tag{7.1}$$

In the case the basis $\{\mathbf{e}^{(1)}(\mathbf{k}), \mathbf{e}^{(2)}(\mathbf{k})\}$ is convenient to be chosen as the basis $\{\mathbf{e}_1^{(l)}(\mathbf{k}), \mathbf{e}_2^{(l)}(\mathbf{k})\}$.

Before calculating the elements of the coherence matrix it is necessary to introduce some auxiliary variables. Let $\mathbf{r}$ be an observation point, remote from the volume of scattering medium ($|\mathbf{r}| \gg |\mathbf{r}_n|, \quad n = 1, \ldots N$), and $\{\mathbf{e}_1(\mathbf{k}), \mathbf{e}_2(\mathbf{k}), \mathbf{k}\}$ be some given basis. Put

$$U_{\alpha\beta}(\mathbf{r}, \mathbf{k}) = (c/8\pi)\langle(\mathbf{e}_\alpha(\mathbf{k}), \mathbf{E}(\mathbf{r}))(\mathbf{e}_\beta(\mathbf{k}), \mathbf{E}(\mathbf{r}))^*\rangle \tag{7.2}$$

and

$$U_{\alpha\beta}^c(\mathbf{r}, \mathbf{k}_0) = (c/8\pi)(\mathbf{e}_\alpha(\mathbf{k}_0), \mathbf{E}^c(\mathbf{r}))(\mathbf{e}_\beta(\mathbf{k}_0), \mathbf{E}^c(\mathbf{r}))^*. \tag{7.3}$$

Obviously, $U_{\alpha\alpha}(\mathbf{r}, \mathbf{k})$ is the value of radiation flux density of polarization state $\mathbf{e}_\alpha$, scattered by the medium in the direction $\mathbf{k}$. Now one should substitute the expression (6.1) for $\mathbf{E}(\mathbf{r})$ into the Eq. (7.2) and perform the averaging. Due to the condition (6.2) one can obtain the result:

$$U_{\alpha\beta}(\mathbf{r}, \mathbf{k}) = \sum_{\alpha',\beta'}^{2} (\mathbf{e}_\alpha(\mathbf{k}), \mathbf{e}^{(\alpha')}(\mathbf{k}_0))(\mathbf{e}_\beta(\mathbf{k}), \mathbf{e}^{(\beta')}(\mathbf{k}_0))^* U_{\alpha'\beta'}^c(\mathbf{r}, \mathbf{k}_0)$$

$$+ \sum_{m,n}^{N}\sum_{\alpha',\beta'}^{2}\langle(\mathbf{e}_\alpha(\mathbf{k}), \mathbf{e}^{(\alpha')}(\mathbf{k}_{n0}))(\mathbf{e}_\beta(\mathbf{k}), \mathbf{e}^{(\beta')}(\mathbf{k}_{m0}))^* F_{\alpha'}(\mathbf{r}, \mathbf{k}_{n0})F_{\beta'}(\mathbf{r}, \mathbf{k}_{m0})\rangle. \tag{7.4}$$

By using the condition (6.2) it is possible to represent the averaged second term in the Eq. (7.4) in the form

$$(c/8\pi)\int_V\int_{s_{\min}}^{s_{\max}} d\mathbf{r}_n ds_n \sum_{\alpha',\beta'}^{2} (\mathbf{e}_\alpha(\mathbf{k}), \mathbf{e}^{(\alpha')}(\mathbf{k}_{n0}))(\mathbf{e}_\beta(\mathbf{k}), \mathbf{e}^{(\beta')}(\mathbf{k}_{n0}))^*\{\langle F_{\alpha'}^{(\circ)}(\mathbf{r}, \mathbf{k}_{n0})F_{\beta'}^{(\circ)*}(\mathbf{r}, \mathbf{k}_{n0})\rangle_n$$

$$+ \int_V\int_{s_{\min}}^{s_{\max}} d\mathbf{r}_m ds_m \rho(\mathbf{r}_m)q(s_m)g_2(\mathbf{r}_n, \mathbf{r}_m; s_n, s_m)\langle F_{\alpha'}^{(\circ)}(\mathbf{r}, \mathbf{k}_{n0})F_{\beta'}^{(\circ)*}(\mathbf{r}, \mathbf{k}_{n0})\rangle_{mn}\} \tag{7.5}$$

Really, in the Eq. (7.5) the term containing the factor $\rho(\mathbf{r}_n)\rho(\mathbf{r}_m)q(s_n)q(s_m)$ (that would correspond to the contribution into $U_{\alpha\beta}(\mathbf{r}, \mathbf{k})$ from coherently scattered radiation) vanishes because all the contributions from coherent scattering have been already taken into account via the term $U_{\alpha\beta}^c(\mathbf{r}, \mathbf{k})$. Besides, due to the conditions

$|\mathbf{r} - \mathbf{r}_n| = O(l), \quad |\mathbf{r}_n - \mathbf{r}_m| = O(R_c)$, one can put $\mathbf{e}^{(\alpha)}(\mathbf{k}_{n0}) \simeq \mathbf{e}^{(\alpha)}(\mathbf{k}_{m0})$. Define further the value

$$
J_{\alpha\beta}(\mathbf{r}, \mathbf{r}', \mathbf{k}') \equiv_{df} \frac{c}{8\pi} \rho(\mathbf{r}') \int\limits_{s_{\min}}^{s_{\max}} q(s') ds' \{ \left\langle F_\alpha^{(\circ)*}(\mathbf{r}, \mathbf{k}'_{n0}) F_\beta^{(\circ)*}(\mathbf{r}, \mathbf{k}'_{n0}) \right\rangle_n
$$

$$
+ \int \int\limits_{s_{\min}}^{s_{\max}} \rho(\mathbf{r}_m) q(s_m) g_2(\mathbf{r}_n, \mathbf{r}_m; s_n, s_m) \delta\left[\frac{\mathbf{r}_n + \mathbf{r}_m}{2} - \mathbf{r}' \right]
$$

$$
\left\langle F_\alpha^{(\circ)}(\mathbf{r}, \mathbf{k}_{n0}) F_\beta^{(\circ)*}(\mathbf{r}, \mathbf{k}_{m0}) \right\rangle_{mn} \}. \tag{7.6}
$$

Substitute now Eq. (7.6) into Eq. (7.5) and after that the Eq. (7.5) into the Eq. (7.4). Then, assuming that $\mathbf{e}^{(\alpha)}(\mathbf{k}_{n0}) \cong \mathbf{e}^{(\alpha)}(\mathbf{k}')$, one can obtain the following expression for $U_{\alpha\beta}(\mathbf{r}, \mathbf{k})$:

$$
U_{\alpha\beta}(\mathbf{r}, \mathbf{k}) = \sum_{\alpha'\beta'}^2 \left(\mathbf{e}_\alpha, \mathbf{e}^{(\alpha')}(\mathbf{k}_0) \right) \left(\mathbf{e}_\beta, \mathbf{e}^{(\beta')}(\mathbf{k}_0) \right)^* U_{\alpha\beta}^c(\mathbf{r}, \mathbf{k}_0)
$$

$$
+ \int_V d\mathbf{r}' \sum_{\alpha'\beta'}^2 \left(\mathbf{e}_\alpha, \mathbf{e}^{(\alpha')}(\mathbf{k}') \right) \left(\mathbf{e}_\beta, \mathbf{e}^{(\beta')}(\mathbf{k}') \right)^* J_{\alpha'\beta'}(\mathbf{r}, \mathbf{r}', \mathbf{k}') \tag{7.7}
$$

At $\alpha = \beta$ the integrand in the second term of the Eq. (7.7) represents the mean of radiation energy density of polarization state $\mathbf{e}_\alpha$, coming to the point $\mathbf{r}$ from unit volume of the scattering medium located in the vicinity of the point $\mathbf{r}'$. Let now the unit vector $\mathbf{k}$ be the vector defining some direction. The coherence matrix elements that are of interest for the considered problem can be defined by the formula (Watson 1969):

$$
J_{\alpha\beta}(\mathbf{r}, \mathbf{k}) = U_{\alpha\beta}^c(\mathbf{r}, \mathbf{k}_0) \delta(\mathbf{k} - \mathbf{k}_0) + \int\limits_{-\mathbf{k}} |\mathbf{r} - \mathbf{r}'|^2 d|\mathbf{r} - \mathbf{r}'| J_{\alpha\beta}(\mathbf{r}, \mathbf{r}'; \mathbf{k}). \tag{7.8}
$$

The integral in the Eq. (7.8) is calculated along the line specified by the vector $-\mathbf{k} = -(\mathbf{r} - \mathbf{r}')(|\mathbf{r} - \mathbf{r}'|)^{-1}$. Comparing the Eq. (7.8) and the Eq. (7.7) one can easily see that

$$
U_{\alpha\beta}(\mathbf{r}, \mathbf{k}) = \int d\Omega_{\mathbf{k}'} \sum_{\alpha'\beta'}^2 \left(\mathbf{e}_\alpha, \mathbf{e}^{(\alpha')}(\mathbf{k}') \right) \left(\mathbf{e}_\beta, \mathbf{e}^{(\beta')}(\mathbf{k}') \right)^* J_{\alpha'\beta'}(\mathbf{r}, \mathbf{k}'). \tag{7.9}
$$

Now all the necessary values for the radiation transport equation have been already defined. However, before the obtaining of the equation itself it is appropriate to specify the approximations for the Green function and the polarization characteristics of coherently scattered radiation that will figure in the process of configuration averaging. Since the locations of the points $\mathbf{r}_n$. $\mathbf{r}_m$, $\mathbf{r}$, $\mathbf{r}'$ satisfy the conditions

$$|\mathbf{r}_n - \mathbf{r}_m| = O(R_c), \quad |\mathbf{r} - \mathbf{r}'| = O(R_c), \quad |\mathbf{r} - \mathbf{r}_n| = O(l), |\mathbf{r} - \mathbf{r}_m| = O(l),$$

$$|\mathbf{r}' - \mathbf{r}_n| = O(l), \quad |\mathbf{r}' - \mathbf{r}_m| = O(l),$$

it is possible to use the approximations

$$G^{(\alpha)}(|\mathbf{r} - \mathbf{r}_m|, \mathbf{k}_{m0}) \simeq e^{i\frac{\omega}{c} n^{(\alpha)}(\mathbf{k}_{n0})\mathbf{k}_{n0}(\mathbf{r}_n - \mathbf{r}_m)} G^{(\alpha)}(|\mathbf{r} - \mathbf{r}_n|; \mathbf{k}_{n0}), \tag{7.10}$$

$$G^{(\alpha)}(|\mathbf{r}' - \mathbf{r}_n|, \mathbf{k}_{n0}) \simeq e^{i\frac{\omega}{c} n^{(\alpha)}(\mathbf{k}_{n0})\mathbf{k}_{n0}(\mathbf{r}'-\mathbf{r})} G^{(\alpha)}(|\mathbf{r} - \mathbf{r}_n|; \mathbf{k}_{n0}), \tag{7.11}$$

$$F_\alpha(\mathbf{r}', \mathbf{r}_n) \simeq e^{i\frac{\omega}{c} n^{(\alpha)}(\mathbf{k}_{n0})(\mathbf{r}'-\mathbf{r})} F_\alpha(\mathbf{r}, \mathbf{r}_n), \tag{7.12}$$

$$F_\alpha(\mathbf{r}, \mathbf{r}_n) \simeq G^{(\alpha)}(|\mathbf{r} - \mathbf{r}_n|, \mathbf{k}_{n0})\{e^{i\frac{\omega}{c} n^{(\alpha)}(\mathbf{k}_{n0})(\mathbf{k}_{n0}-\mathbf{k}_0)(\mathbf{r}_n - \mathbf{r}_m)}$$

$$\cdot \sum_\beta^2 A_{\alpha\beta}^{(\circ)}(\mathbf{r}_n, s_n; \mathbf{k}_0 \to \mathbf{k}_{n0}) E_\beta^c(\mathbf{r}_n, \mathbf{k}_0)$$

$$+ \sum_{p \neq n} e^{i\frac{\omega}{c} n^{(\alpha)}(\mathbf{k}_{n0})(\mathbf{k}_{n0}-\mathbf{k}_{mn})(\mathbf{r}_n - \mathbf{r}_p)} \sum_{\beta'}^2 A_{\alpha\beta'}^{(\circ)}(\mathbf{r}_n, s_n; \mathbf{k}_{np} - \mathbf{k}_{n0}) F_{\beta'}^{(\circ)}(\mathbf{r}_n, \mathbf{r}_p)\}. \tag{7.13}$$

Substitute now the expression for $F_\alpha(\mathbf{r}', \mathbf{r}_n)$ into the formula (7.6) and perform the configuration averaging operation for the Eq. (7.6) under accounting the Eqs. (7.10)–(7.13). Using the vector notations we obtain the following expression for the value $\mathbf{J}(\mathbf{r}, \mathbf{r}', \mathbf{k})$:

$$\mathbf{J}(\mathbf{r}, \mathbf{r}', \mathbf{k}) \equiv_{df} \frac{c}{8\pi} \int_{s_{min}}^{s_{max}} \rho(\mathbf{r}_n, s_n) ds_n \{\langle \mathbf{F}^{(\circ)}(\mathbf{r}, \mathbf{k}_{n0}) \otimes \mathbf{F}^{(\circ)*}(\mathbf{r}, \mathbf{k}_{n0})\rangle_n +$$

$$+ \int_V \int_{s_{min}}^{s_{max}} \rho(\mathbf{r}_m, s_m) g_2(\mathbf{r}_n, \mathbf{r}_m; s_n, s_n) d\mathbf{r}' ds_n \delta\left[\frac{\mathbf{r}_n + \mathbf{r}_m}{2} - \mathbf{r}'\right] \langle \mathbf{F}^{(\circ)}(\mathbf{r}, \mathbf{k}_{n0})$$

$$\otimes \mathbf{F}^{(\circ)*}(\mathbf{r}, \mathbf{k}_{n0})\rangle_{mn}\} \tag{7.14}$$

Finally we have:

$$\mathbf{J}(\mathbf{r}, \mathbf{r}'; \mathbf{k}) = [\hat{G}^{(\circ)}(|\mathbf{r} - \mathbf{r}'|, \mathbf{k}) \otimes \hat{G}^{(\circ)*}(|\mathbf{r} - \mathbf{r}'|, \mathbf{k})]\{\hat{P}(\mathbf{r}'; \mathbf{k}_0 \to \mathbf{k})\mathbf{J}^c(\mathbf{r}', \mathbf{k}_0)$$

$$+ \int d\Omega_{\mathbf{k}'} \hat{P}(\mathbf{r}'; \mathbf{k}' \to \mathbf{k})\mathbf{J}(\mathbf{r}', \mathbf{k}')\}, \tag{7.15}$$

where

$$\hat{P}(\mathbf{r}; \mathbf{k}' \to \mathbf{k}) = \hat{L}(\mathbf{k})\hat{\Gamma}(\mathbf{r}, \mathbf{k}' \to \mathbf{k})\hat{L}^{-1}(\mathbf{k}'), \tag{7.16}$$

$$\hat{L}(\mathbf{k}) = \hat{L}(\mathbf{k}) \otimes \hat{L}^*(\mathbf{k}), \tag{7.17}$$

$$\hat{\Gamma}(\mathbf{r}, \mathbf{k}' \to \mathbf{k}) = \int_{S_{\min}}^{S_{\max}} ds \rho(\mathbf{r}, s) \left\{ \hat{A}^{(s)}(\mathbf{r}, s; \mathbf{k}' \to \mathbf{k}) \otimes \hat{A}^{(s)*}(\mathbf{r}, s; \mathbf{k}' \to \mathbf{k}) \right\} +$$

$$\int_V \int_{S_{\min}}^{S_{\max}} d\mathbf{r}' ds' \rho(\mathbf{r}', s') g_2(\mathbf{r}, \mathbf{r}'; s, s') e^{i\hat{n}(\mathbf{k})(\mathbf{k} \sim \mathbf{k}')(\mathbf{r} \sim \mathbf{r}')}.$$

$$\cdot \left\{ \hat{A}^{(s)}(\mathbf{r}', s'; \mathbf{k}' \to \mathbf{k}) \otimes \hat{A}^{(s)*}(\mathbf{r}', s'; \mathbf{k}' \to \mathbf{k}) \right\}, \tag{7.18}$$

$$e^{i\hat{n}(\mathbf{k})x} \equiv_{df} \begin{pmatrix} e^{in^{(1)}(\mathbf{k})x} & 0 \\ 0 & e^{in^{(2)}(\mathbf{k})x} \end{pmatrix} \otimes \begin{pmatrix} e^{in^{(1)}(\mathbf{k})x} & 0 \\ 0 & e^{in^{(2)}(\mathbf{k})x} \end{pmatrix}. \tag{7.19}$$

Multiplying both the sides of Eq. (7.14) by $\left| \mathbf{r} - \mathbf{r}' \right|^2 d \left| \mathbf{r} - \mathbf{r}' \right|$ and performing the integration along the direction $-\mathbf{k}$, one can obtain the equation

$$\mathbf{J}(\mathbf{r}, \mathbf{k}) = \mathbf{J}^c(\mathbf{r}, \mathbf{k}_0) \delta(\mathbf{k} - \mathbf{k}_0) +$$

$$\int_{-\mathbf{k}} \left| \mathbf{r} - \mathbf{r}' \right|^2 d \left| \mathbf{r} - \mathbf{r}' \right| \hat{G}^{(\circ)}(\left| \mathbf{r} - \mathbf{r}' \right|, \mathbf{k}) \otimes \hat{G}^{(\circ)*}(\left| \mathbf{r} - \mathbf{r}' \right|, \mathbf{k})$$

$$\int d\Omega_{\mathbf{k}'} \hat{P}(\mathbf{r}'; \mathbf{k}' \to \mathbf{k}) \mathbf{J}(\mathbf{r}', \mathbf{k}') \tag{7.20}$$

At last, using the expressions for refractive indices of normal waves in anisotropic medium and attracting the optical theorem according to which the relations take place

$$\sigma_t^{(\alpha)}(\mathbf{k}) = 4\pi \left(\frac{\omega}{c} \right)^{-1} \operatorname{Im} A^{(\alpha)}(\mathbf{k} \to \mathbf{k}), \tag{7.21}$$

the following expression for $\hat{G}^{(\circ)}(\left| \mathbf{r} - \mathbf{r}' \right|, \mathbf{k}) \otimes \hat{G}^{(\circ)*}(\left| \mathbf{r} - \mathbf{r}' \right|, \mathbf{k})$ can be obtained

$$\hat{G}^{(\circ)}(r, \mathbf{k}) \otimes \hat{G}^{(\circ)*}(r, \mathbf{k})$$

$$= r^{-2} \begin{bmatrix} e^{-\sigma^{(1)}(\mathbf{k})r} & 0 & 0 & 0 \\ 0 & e^{i\frac{\omega}{c}[\tilde{n}^{(1)}(\mathbf{k}) - \tilde{n}^{(2)}(\mathbf{k})] - \frac{\sigma^{(1)} + \sigma^{(2)}}{2} r} & 0 & 0 \\ 0 & 0 & e^{i\frac{\omega}{c}[\tilde{n}^{(1)}(\mathbf{k}) - \tilde{n}^{(2)}(\mathbf{k})] - \frac{\sigma^{(1)} + \sigma^{(2)}}{2} r} & 0 \\ 0 & 0 & 0 & e^{-\sigma^{(2)}(\mathbf{k})r} \end{bmatrix} \tag{7.22}$$

Now we are able to write the final version of integral transport equation for polarization vector $\mathbf{J}(\mathbf{r}, \mathbf{k})$, calculated in the basis $\{\mathbf{e}^{(1)}(\mathbf{k}), \mathbf{e}^{(2)}(\mathbf{k})\}$:

$$\mathbf{J}(\mathbf{r}, \mathbf{k}) = \mathbf{J}^c(\mathbf{r}, \mathbf{k}_0) \delta(\mathbf{k} - \mathbf{k}_0) + \int_{-\mathbf{k}} d \left| \mathbf{r} - \mathbf{r}' \right| e^{\hat{\mathbf{D}}(\left| \mathbf{r} - \mathbf{r}' \right|)} \int d\Omega_{\mathbf{k}'} \hat{P}(\mathbf{r}'; \mathbf{k}' \to \mathbf{k}),$$

$$\tag{7.23}$$

$$e^{\hat{\mathbf{D}}(r)} \equiv_{df} r^2 \Big[\hat{G}^{(\circ)}(r, \mathbf{k}) \otimes \hat{G}^{(\circ)*}(r, \mathbf{k})\Big]. \tag{7.24}$$

8 Integro-Differential Radiation Transport Equation. The Drift Operator

By differentiation of the Eq. (7.23) along the direction $\mathbf{k}$ one can obtain the integro-differential radiation transport equation that can be written as

$$(\mathbf{k}, \nabla)\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k}) + \hat{D}^{(0)}(\mathbf{k})\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k}) = \int \hat{P}^{(0)}(\mathbf{r}; \mathbf{k}' \to \mathbf{k})\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k}')d\Omega_{\mathbf{k}'}, \tag{8.1}$$

$\hat{D}^{(0)}(\mathbf{k}) =$

$$\begin{bmatrix} \sigma_t^{(1)}(\mathbf{k}) & 0 & 0 & 0 \\ 0 & \frac{1}{2}\big[\sigma^{(1)}(\mathbf{k}) + \sigma^{(2)}(\mathbf{k})\big] + \mathrm{i}\frac{\omega}{c}\big[\tilde{n}^{(1)} - \tilde{n}^{(2)}\big] & 0 & 0 \\ 0 & 0 & \frac{1}{2}\big[\sigma^{(1)}(\mathbf{k}) + \sigma^{(2)}(\mathbf{k})\big] - \mathrm{i}\frac{\omega}{c}\big[\tilde{n}^{(1)} - \tilde{n}^{(2)}\big] & 0 \\ 0 & 0 & 0 & \sigma_t^{(2)}(\mathbf{k}) \end{bmatrix} \tag{8.2}$$

The Eqs. (8.1)–(8.2) looks like vector radiation transport equation for optically isotropic medium with the following exceptions:

(1) the Eq. (8.1) is written for the polarization vector $\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k})$ defined in the basis of eigenvectors of the operator $\hat{A}(\mathbf{r}, \mathbf{k} \to \mathbf{k})$;

(2) the differential operator of integro-differential radiation transport equation for optically anisotropic medium differs from that one for optically isotropic medium by the presence of matrix operator $\hat{D}(\mathbf{k})$ that is responsible for interference effects caused by the different phase velocities and absorptions of normal waves in anisotropic medium;

(3) the structure for scattering operator $\hat{P}(\mathbf{r}; \mathbf{k}', \mathbf{k})$ (see the Eqs. (7.16)–(7.19)) reveals the relation between the scattering properties of medium volume element and microscopic medium characteristics (the amplitude scattering operator of individual scatterers) and the features of scatterer statistical distribution.

The Eq. (8.1) can be transformed into that one written for another polarization vector $\mathbf{I}(\mathbf{r}, \mathbf{k})$, connected with $\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k})$ by linear transformation and defined in some basis $\{\mathbf{e}_1^{(l)}(\mathbf{k}), \ \mathbf{e}_2^{(l)}(\mathbf{k})\}$, naturally related to geometry of scattering medium. In particular, the transport equation can be written for the Stokes vector, $\mathbf{I}^S(\mathbf{r}, \mathbf{k})$. Let $\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k})$ be the polarization vector defined in the basis $\{\mathbf{e}^{(1)}(\mathbf{k}), \ \mathbf{e}^{(2)}(\mathbf{k})\}$ of normal waves (in general, not orthonormal), the polarization vector $\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k})$ be the same polarization vector defined in some "laboratory" orthonormal basis $\{\mathbf{e}_1^{(l)}(\mathbf{k}), \ \mathbf{e}_2^{(l)}(\mathbf{k})\}$, and $\hat{\Lambda}^0(\mathbf{k})$ be the matrix with elements

$$\Lambda_{\alpha\beta}^0(\mathbf{k}) = (\mathbf{e}^{(\alpha)}(\mathbf{k}), \mathbf{e}^{(\beta)}(\mathbf{k})), \quad \alpha, \beta = 1, 2.$$

Then, as one can be convinced, there is the following relation between $\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k})$ and $\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k})$:

$$\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k}) = \hat{L}(\mathbf{k})\mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k}), \ \mathbf{J}^{(0)}(\mathbf{r}, \mathbf{k}) = \hat{\Lambda}(\mathbf{k})\hat{L}^+(\mathbf{k})\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k}), \qquad (8.3)$$

where

$$\hat{L}(\mathbf{k}) = \hat{\mathcal{L}}^T(\mathbf{k}) \otimes (\hat{\mathcal{L}}^T(\mathbf{k}))^*, \ \hat{\mathcal{L}}_{\alpha\beta} = (\mathbf{e}^{(\alpha)}, \mathbf{e}_\beta^{(l)}), \quad \alpha, \beta = 1, 2, \qquad (8.4)$$

$$\hat{\Lambda}(\mathbf{k}) = \left[\hat{\Lambda}_0^T(\mathbf{k})\right]^{-1} \otimes \left[\hat{\Lambda}_0(\mathbf{k})\right]^{-1} \qquad (8.5)$$

$$\hat{\Lambda}_0(\mathbf{k}) = \begin{bmatrix} 1 & a \\ a^* & 1 \end{bmatrix}, \quad a = (\mathbf{e}^{(1)}, \mathbf{e}^{(2)}), \qquad (8.6)$$

$$\hat{\Lambda}(\mathbf{k}) = \frac{1}{1 - aa^*} \begin{bmatrix} 1 & -a & -a^* & aa^* \\ -a^* & 1 & a^2 & -a^* \\ -a & a^2 & 1 & -a \\ aa^* & -a & -a^* & 1 \end{bmatrix} \qquad (8.7)$$

Apply now the linear transformation ΛL^+ to the transport Eq. (8.1) taking into account the relations (8.3). As a result we obtain the following transport equation for the polarization vector $\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k})$:

$$(\mathbf{k}, \nabla)\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k}) + \hat{D}^{(l)}(\mathbf{k})\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k}) = \int \hat{P}^{(l)}(\mathbf{r}; \mathbf{k}' \to \mathbf{k})\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k}')d\Omega_{\mathbf{k}'}, \qquad (8.8)$$

where

$$\hat{D}^{(l)}(\mathbf{k}) = \hat{L}(\mathbf{k}) \, \hat{D}^{(0)}(\mathbf{k}) \, \hat{\Lambda}(\mathbf{k}) \, \hat{L}^+(\mathbf{k}),$$
$$\hat{P}^{(l)}(\mathbf{r}, \mathbf{k}' \to \mathbf{k}) = \hat{L}\hat{P}^{(0)}(\mathbf{r}, \mathbf{k}' \to \mathbf{k})\hat{\Lambda}\hat{L}^+. \qquad (8.9)$$

If it is necessary to write the transport equation for another polarization vector, defined in the same basis, $\mathbf{I}(\mathbf{r}, \mathbf{k}) = \hat{T}\mathbf{J}^{(l)}(\mathbf{r}, \mathbf{k})$, it can be realized via the following transformation the transport equation:

$$(\mathbf{k}, \nabla)\mathbf{I}(\mathbf{r}, \mathbf{k}) + \hat{D}(\mathbf{k})\mathbf{I}(\mathbf{r}, \mathbf{k}) = \int \hat{P}(\mathbf{r}; \mathbf{k}' \to \mathbf{k})\mathbf{I}(\mathbf{r}, \mathbf{k}')d\Omega_{\mathbf{k}'}, \qquad (8.10)$$

where

$$\hat{D}(\mathbf{k}) = \hat{\tilde{L}} \, \hat{\tilde{D}}(\mathbf{k})\hat{\tilde{\Lambda}} \, \hat{\tilde{L}}^+, \quad \hat{P}(\mathbf{r}; \mathbf{k}' \to \mathbf{k}) = \hat{\tilde{L}}(\mathbf{k})\hat{\tilde{\Gamma}}(\mathbf{r}; \mathbf{k}' \to \mathbf{k})\tilde{L}^{-1}(\mathbf{k}), \qquad (8.11)$$

$$\hat{\tilde{A}} \equiv_{df} \hat{T} \, \hat{A} \, \hat{T}^{-1}.$$

In particular, for the Stokes polarization vector $(\mathbf{I}(\mathbf{r}, \mathbf{k}) = \mathbf{I}^S(\mathbf{r}, \mathbf{k}))$ we have

$$\hat{T}^S = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix}, \quad (\hat{T}^S)^{-1} = 0.5 \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & i \\ 1 & -1 & 0 & 0 \end{bmatrix} \quad (8.12)$$

Now we can write the radiation transport equation for the Stokes vector for some typical particular situations. First of all notice that case of homogeneous optically isotropic medium is realized under the condition

$$\hat{\mathcal{A}}(\mathbf{r}, \mathbf{k} \to \mathbf{k}) = \hat{\mathcal{A}}(\mathbf{k} \to \mathbf{k}) = a_0 \hat{I}_2, \quad (8.13)$$

where a_0 is a constant, and $\hat{I}_2$ is unit matrix of the second order. In the case any orthonormal basis can be chosen as the basis $\{\mathbf{e}^{(1)}(\mathbf{k}), \mathbf{e}^{(2)}(\mathbf{k})\}$, the basis vectors, located in the plane orthogonal to the vector $\mathbf{k}$, being either real-valued or complex-valued. If radiation transport problem is considered in the slab, $\mathbf{n}_0$ being the normal to the slab surface, it is convenient to choose

$$\mathbf{e}_2^{(l)}(\mathbf{k}) = \mathbf{e}_2^{(0)}(\mathbf{k}) = \frac{\mathbf{n}_0 \times \mathbf{k}}{|\mathbf{n}_0 \times \mathbf{k}|}, \quad \mathbf{e}_1^{(l)}(\mathbf{k}) = \mathbf{e}_1^{(0)}(\mathbf{k}) = \mathbf{e}_2^{(0)} \times \mathbf{k}. \quad (8.14)$$

So, in the case of isotropic medium we have

$$\hat{L}(\mathbf{k}) = \hat{\Lambda}(\mathbf{k}) = \hat{I}_4, \quad (8.15)$$

where $\hat{I}_4$ is the unit diagonal matrix of fourth order. Therefore, the matrices of the drift operator $\hat{D}(\mathbf{k})$ and of the scattering operator $\hat{P}$ of transport equation can be written in the form

$$\hat{D}(\mathbf{k}) = \sigma_t \hat{I}_4, \quad \sigma_t = \operatorname{Im} a_0, \quad (8.16)$$

$$\hat{P}(\mathbf{k}' \to \mathbf{k}) = \hat{L}^S(\mathbf{k}) \hat{\Gamma}^S(\mathbf{k}' \to \mathbf{k}) [L^S(\mathbf{k}')]^{-1}, \quad (8.17)$$

$$\hat{\Gamma}^S(\mathbf{k}' \to \mathbf{k}) = \hat{T}^S \hat{\Gamma}(\mathbf{k}' \to \mathbf{k}) [\hat{T}^S]^{-1}, \quad (8.18)$$

$$\hat{L}^S(\mathbf{k}) = \hat{T}^S [\hat{\mathcal{L}}^T(\mathbf{k}) \otimes (\hat{\mathcal{L}}^T(\mathbf{k}))^*] [\hat{T}^S]^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & a^2 - b^2 & 2ab & 0 \\ 0 & -2ab & a^2 - b^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (8.19)$$

$$[\hat{L}^S(\mathbf{k}')]^{-1} = \hat{T}^S[\hat{\mathcal{L}}^*(\mathbf{k}') \otimes (\hat{\mathcal{L}}^T(\mathbf{k}'))^*][\hat{T}^S]^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & a'^2 - b'^2 & -2a'b' & 0 \\ 0 & 2a'b' & a'^2 - b'^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$(8.20)$$

$$\hat{\mathcal{L}}_{\alpha\beta}(\mathbf{k}) = (\mathbf{e}_\alpha^{(l)}(\mathbf{k}), \mathbf{e}_\beta^{(s)}(\mathbf{k})), \quad \alpha, \beta = 1, 2, \tag{8.21}$$

$$a = \frac{-(\mathbf{n}_0, \mathbf{k}') + (\mathbf{n}_0, \mathbf{k})(\mathbf{k}, \mathbf{k}')}{|\mathbf{k} \times \mathbf{k}'||\mathbf{n}_0 \times \mathbf{k}|}, \quad a' = \frac{(\mathbf{n}_0, \mathbf{k}) - (\mathbf{n}_0, \mathbf{k}')(\mathbf{k}, \mathbf{k}')}{|\mathbf{k} \times \mathbf{k}'||\mathbf{n}_0 \times \mathbf{k}'|},$$

$$b = \frac{\mathbf{n}_0 \mathbf{k}\,\mathbf{k}'}{|\mathbf{k} \times \mathbf{k}'||\mathbf{n}_0 \times \mathbf{k}|}, \quad b' = \frac{\mathbf{n}_0 \mathbf{k}\,\mathbf{k}'}{|\mathbf{k} \times \mathbf{k}'||\mathbf{n}_0 \times \mathbf{k}'|}. \tag{8.22}$$

The matrix $\hat{\Gamma}$ in the formula (8.18) is defined by the expression (7.18), and the vectors $\mathbf{e}_1^{(s)}(\mathbf{k})$, $\mathbf{e}_2^{(s)}(\mathbf{k})$—by the formulas (3.10), (3.11).

Thus, the vector transport Eq. (8.10) with the operators $\hat{D}$ and $\hat{P}$, defined by formulas (8.16) and (8.17), describing transfer of multiply scattered electromagnetic radiation in homogeneous optically isotropic medium, coincides with transport equation that can be phenomenologically derived based on consideration of energy balance relations (see, for example, (Chandrasekhar 1950)).

The transport equation for slab of optically anisotropic medium with linear birefringence (linear double refraction) can deliver another example. Let the normal $\mathbf{n}_0$ to the slab surface coincides with the medium optical axis. As it is known, in the medium two linearly polarized electromagnetic waves can propagate in each direction $\mathbf{k}$ in the medium, their refractive indices $\tilde{n}^{(1)}(\mathbf{k})$, $\tilde{n}^{(2)}(\mathbf{k})$ and polarization vectors $\mathbf{e}_1^{(0)}(\mathbf{k})$, $\mathbf{e}_2^{(0)}(\mathbf{k})$ being specified by the formulas (8.14). So, the matrix $\hat{\mathcal{A}}(\mathbf{k} \to \mathbf{k})$ is diagonal in the basis $\{\mathbf{e}_1^{(0)}(\mathbf{k}), \mathbf{e}_2^{(0)}(\mathbf{k})\}$,

$$\hat{\mathcal{A}}(\mathbf{k} \to \mathbf{k}) = \begin{bmatrix} \mathcal{A}^{(1)}(\mathbf{k}) & 0 \\ 0 & \mathcal{A}^{(2)}(\mathbf{k}) \end{bmatrix}, \tag{8.23}$$

and therefore, just as in the case optically isotropic medium, the relationship (8.15) is valid. It means that the matrix of operator $\hat{P}(\mathbf{k}' \to \mathbf{k})$ can be obtained according to the formulas (8.19)–(8.22). The matrix of operator $\hat{D}(\mathbf{k})$ can be also easily calculated and written as

$$\hat{D}(\mathbf{k}) = \hat{T}^S D^{(l)}[\hat{T}^S]^{-1}$$

$$= \begin{bmatrix} 0.5[\sigma^{(1)}(\mathbf{k}) + \sigma^{(2)}(\mathbf{k})] & 0.5[\sigma^{(1)}(\mathbf{k}) - \sigma^{(2)}(\mathbf{k})] & 0 & 0 \\ 0.5[\sigma^{(1)}(\mathbf{k}) - \sigma^{(2)}(\mathbf{k})] & 0.5[\sigma^{(1)}(\mathbf{k}) + \sigma^{(2)}(\mathbf{k})] & 0 & 0 \\ 0 & 0 & 0.5[\sigma^{(1)}(\mathbf{k}) + \sigma^{(2)}(\mathbf{k})] & -(\omega/c)[n^{(1)} - n^{(2)}] \\ 0 & 0 & (\omega/c)[n^{(1)} - n^{(2)}] & 0.5[\sigma^{(1)}(\mathbf{k}) + \sigma^{(2)}(\mathbf{k})] \end{bmatrix}, \tag{8.24}$$

The drift operator $\hat{D}(\mathbf{k})$ defines the law of propagation of coherently scattered radiation in optically anisotropic medium. Indeed, if non-coherent scattering is absent $(\hat{P}(\mathbf{k'} \rightarrow \mathbf{k}) \equiv 0)$, then from Eq. (7.20) one obtains

$$\mathbf{I}(\mathbf{r}, \mathbf{k}) = \mathbf{I}^c(\mathbf{r}, \mathbf{k}), \tag{8.25}$$

the polarization vector of coherently scattered radiation being satisfied to the differential equation

$$(\mathbf{k}, \nabla)\mathbf{I}^c(\mathbf{r}, \mathbf{k}) + \hat{D}(\mathbf{k})\mathbf{I}^c(\mathbf{r}, \mathbf{k}) = 0 \tag{8.26}$$

The problem of coherently scattered radiation transfer through optically anisotropic media, that represents an independent interest, was studied in a number of papers (see for example (Sazonov 1969; Zheleznyakov et al. 1974; Sazonov 1974)). The paper (Silant'ev 1971) contains the analysis of drift operator for various anisotropic media composed of oriented isolated scatterers, small compared radiation wavelength, and the analysis of properties of radiation coherently scattered and transferred in slabs of the anisotropic media.

In the simplest case of anisotropic medium with linear birefringence the analysis of drift operator $\hat{D}(\mathbf{k})$ allows to extract some features of coherently scattered radiation by a slab of the anisotropic medium. For example, one can obtain the result: at oblique incidence of parallel linearly polarized radiation flux to the slab surface of the anisotropic medium the coherently scattered and transferred by the slab radiation acquires elliptical polarization.

Supplement

Eigenvalues and eigenvectors of the operator $\hat{\mathcal{A}}(\mathbf{k} \rightarrow \mathbf{k})$

Consider first the case of Hermitian matrix $\hat{\mathcal{A}}(\mathbf{k} \rightarrow \mathbf{k})$, representing it as

$$\hat{\mathcal{A}}(\mathbf{k} \rightarrow \mathbf{k}) = \begin{bmatrix} \mathcal{A}'_{11} & \mathcal{A}'_{12} \\ \mathcal{A}'_{12} & \mathcal{A}'_{22} \end{bmatrix} + i \begin{bmatrix} 0 & \mathcal{A}''_{12} \\ -\mathcal{A}''_{12} & 0 \end{bmatrix}, \tag{S.1}$$

where

$$\mathcal{A}'_{\alpha\beta} = \operatorname{Re} \mathcal{A}_{\alpha\beta}, \quad \mathcal{A}''_{\alpha\beta} = \operatorname{Im} \mathcal{A}_{\alpha\beta}.$$

The eigenvalues $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ of a Hermitian matrix are real values and are the roots of the characteristic polynomial

$$\lambda^2 - t\lambda + d,$$

where

$$t = \operatorname{Tr} \hat{A}, \quad d = \det \hat{A}.$$

Put at first $\mathcal{A}''_{12} = 0$ (what means from physical viewpoint that non-absorbing non-magnetic anisotropic medium is considered). Then we have the following eigenvalues and and orthonormal eigenvectors of the operator $\hat{A}(\mathbf{k} \to \mathbf{k})$

$$\mathcal{A}^{(1,2)} = \frac{1}{2}\left\{(\mathcal{A}'_{11} + \mathcal{A}'_{22}) \pm \sqrt{(\mathcal{A}'_{11} - \mathcal{A}'_{22})^2 + 4\mathcal{A}'^2_{11}}\right\}, \tag{S.2}$$

$$\mathbf{e}^{(0)(1)} = (C_o^2 + \mathcal{A}'^2_{12})^{-1/2}\begin{bmatrix} \mathcal{A}'_{12} \\ C_o \end{bmatrix}, \tag{S.3}$$

$$\mathbf{e}^{(0)(2)} = (C_o^2 + \mathcal{A}'^2_{12})^{-1/2}\begin{bmatrix} C_o \\ -\mathcal{A}'_{12} \end{bmatrix}. \tag{S.4}$$

where

$$C_0 = \frac{1}{2}\left\{-(\mathcal{A}'_{11} - \mathcal{A}'_{22}) + \sqrt{(\mathcal{A}'_{11} - \mathcal{A}'_{22})^2 + 4\mathcal{A}'^2_{12}}\right\}. \tag{S.5}$$

So, in the anisotropic medium along any direction $\mathbf{k}$ two linearly polarized waves in mutually orthogonal polarization states, defined by vectors $\mathbf{e}^{(0)(1)}$ and $\mathbf{e}^{(0)(2)}$, can propagate without absorption with phase velocities $v_{ph}^{(\alpha)}(\mathbf{k}) = c/\mathcal{A}^{(\alpha)}(\mathbf{k})$, $\alpha = 1, 2$. Let now $\mathcal{A}''_{12} \neq 0$. Then in the basis $\{\mathbf{e}^{(0)(1)}, \mathbf{e}^{(0)(2)}\}$ the operator can be written as

$$\hat{\mathcal{A}}(\mathbf{k} \to \mathbf{k}) = \begin{bmatrix} \mathcal{A}^{(0)1} & i\mathcal{A}''_{12} \\ -i\mathcal{A}''_{12} & \mathcal{A}^{(0)2} \end{bmatrix}. \tag{S.6}$$

The eigenvalues and eigenvectors of $\hat{\mathcal{A}}(\mathbf{k} \to \mathbf{k})$ are now defined by the expressions:

$$\mathcal{A}^{(1,2)} = \frac{1}{2}\{(\mathcal{A}^{(0)1} + \mathcal{A}^{(0)2}) \pm \sqrt{(\mathcal{A}^{(0)1} - \mathcal{A}^{(0)2})^2 + 4\mathcal{A}''^2_{12}}\}, \tag{S.7}$$

$$\mathbf{e}^{(1)} = (C_1^2 + \mathcal{A}''^2_{12})^{-1/2}\begin{bmatrix} \mathcal{A}''_{12} \\ -i C_1 \end{bmatrix}, \quad \mathbf{e}^{(2)} = (C_1^2 + \mathcal{A}''^2_{12})^{-1/2}\begin{bmatrix} C_1 \\ -i \mathcal{A}''_{12} \end{bmatrix}, \tag{S.8}$$

$$C_1 = \frac{1}{2}\{-(\mathcal{A}^{(0)1} - \mathcal{A}^{(0)2}) + \sqrt{(\mathcal{A}^{(0)1} - \mathcal{A}^{(0)2})^2 + 4\mathcal{A}''^2_{12}}\}. \tag{S.9}$$

Therefore in the case of Hermittean and complex-valued $\hat{\mathcal{A}}(\mathbf{k} \to \mathbf{k})$ two elliptically polarized waves in mutually opposite polarization states can propagate in anisotropic medium along any direction $\mathbf{k}$ without absorption. For these waves the

orientations of large axes of polarization ellipse coincide with the directions $\mathbf{e}^{(0)(1)}$ and $\mathbf{e}^{(0)(2)}$. In particular, if $\mathcal{A}''_{12} \ll \mathcal{A}^{(0)1} - \mathcal{A}^{(0)2}$ we have

$$\mathcal{A}^{(1,2)} \simeq \mathcal{A}^{(0)(1,2)} + \frac{\mathcal{A}''_{12}}{\mathcal{A}^{(0)(1)} - \mathcal{A}^{(0)(2)}}, \tag{S.10}$$

$$\mathbf{e}^{(1)} = \left[1 + \left(\frac{\mathcal{A}''_{12}}{\mathcal{A}^{(0)(1)} - \mathcal{A}^{(0)(2)}} \right)^2 \right]^{-1/2} \left[\begin{array}{c} 1 \\ -i \frac{\mathcal{A}''_{12}}{\mathcal{A}^{(0)(1)} - \mathcal{A}^{(0)(2)}} \end{array} \right],$$

$$\mathbf{e}^{(1)} = \left[1 + \left(\frac{\mathcal{A}''_{12}}{\mathcal{A}^{(0)(1)} - \mathcal{A}^{(0)(2)}} \right)^2 \right]^{-1/2} \left[\begin{array}{c} \frac{\mathcal{A}''_{12}}{\mathcal{A}^{(0)(1)} - \mathcal{A}^{(0)(2)}} \\ i \end{array} \right]. \tag{S.11}$$

Therefore, at small value $\mathcal{A}''_{12}$ the normal waves are weakly elliptically polarized.

The case $\mathcal{A}^{(0)(1)} = \mathcal{A}^{(0)(2)}$ corresponds to magneto-optical effect in isotropic medium. Indeed, we have

$$\mathcal{A}^{(1,2)} \simeq \mathcal{A}^{(0)} \pm \mathcal{A}''_{12}, \tag{S.12}$$

$$\mathbf{e}^{(1,2)} = 2^{-1/2} \left[\begin{array}{c} 1 \\ \pm i \end{array} \right], \tag{S.13}$$

and so the circular birefringence takes place in the medium.

In general case the eigenvalues of $\hat{\mathcal{A}}(\mathbf{k} \to \mathbf{k})$ are complex-valued, and the eigenvectors are not orthogonal. The normal matrices, satisfying the condition $\hat{\mathcal{A}} \hat{\mathcal{A}}^+ = \hat{\mathcal{A}}^+ \hat{\mathcal{A}}$, possess orthogonal eigenvectors.

References

Barabanenkov YuN (1969) On the spectral theory of radiation transport equations. JETP 29(4):679

Barabanenkov YuN (1970) Perturbation theory in denominator for average double green function (in Russian). Izv Vyschikh Uch Zav Radiofizika 13(1):106–114

Barabanenkov YuN (1971) On Fraunhofer approximation for multiple wave scattering theory (in Russian). Izv Vyschikh Uch Zav Radiofizika 14(2):234–243

Barabanenkov YuN (2009) Asymptotic limit of radiative transport theory in problems of multiple scattering of waves in random inhomogeneous media. Phys Usp 52(5):502

Barabanenkov YuN, Finkelberg VM (1967) Radiation transport equation for correlated scatterers. JETP 26(3):587

Barabanenkov YuN, Kravtsov YuA, Rytov SM, Tatarsky VI (1971) Status of the theory of propagation of waves in randomly inhomogeneous medium (in Russian). Sov Phys Usp 13(5):551

Barabanenkov YuN, Vinogradov AG, Kravtsov YuA, Tatarsky VI (1972) Usage of theory of wave multiple scattering for derivation of the radiative transfer equation for statistically inhomogeneous medium (in Russian). Izv Vyschikh Uch Zav Radiofizika 15(12):1852

Chandrasekhar S (1950) Radiative transfer. Oxford University Press

Dolginov AZ, Gnedin YuN, Silant'ev NA (1970) Photon polarization and frequency change in multiple scattering. J Quant Spectrosc Rad Trans 10(7):707–754

Dyson FJ (1949) The radiation theories of Tomonaga, Schwinger, and Feynman. Phys Rev 75:486

Foldy LL (1945) The multiple scattering of waves, I. General theory of isotropic scattering by randomly distributed scatterers. Phys Rev 67:107–119

Gnedin YuN, Dolginov AZ (1964) Theory of multiple scattering. JETP 18(3):784

Gnedin YuN, Dolginov AZ (1965) Theory of multiple scattering II. JETP 21(2):364

Gnedin YuN, Dolginov AZ (1967) Radiative transfer in a finite medium. Soviet Astron 10:637

Gnedin YuN, Dolginov AZ, Silant'ev NA (1970a) Intensity and polarization of radiation multiply scattered by freely oriented particles or medium fluctuations. JETP 30(3):540

Gnedin YuN, Dolginov AZ, Silant'ev NA (1970b) (in Russian) JETP 58:319

Gnedin YuN, Dolginov AZ, Silant'ev NA (1970c) (in Russian) JETP 10(7):865

Kruglov VI (1978) To the theory of oscillatory non-equilibrium radiation of two-atom molecules. Dissertation, Minsk (in Russian)

Kuzmina MG (1976) Polarized radiation transport equation in anisotropic media. Preprint of KIAM RAS, 68 (in Russian)

Law CW, Watson KM (1970) Radiation transport along curved ray paths. J Math Phys 11:3125–3137

Lax M (1951) Multiple scattering of waves. Rev Modern Phys 23:287–310

Mandt C, Tsang LD (1992) Backscattering enhancement from a random distribution of large discrete spherical scatterers with a size distribution. J Opt Soc Am A 9(12):2246–2251

Mishchenko MI (2008a) Multiple scattering by particles embedded in an absorbing medium. 2. Radiative transfer equation. J Quant Spect Rad Trans 109:2386–2390

Mishchenko MI (2008b) Multiple scattering, radiative transfer, and weak localization in discrete random media: unified microphysical approach. Rev Geophys 46:1–33

Mishchenko MI, Travis LD, Lacis AA (2002) Scattering, absorption, and emission of light by small particles. Cambridge University Press

Mishchenko MI, Dlugach JM, Yurkin MA, Bi L, Cairns B, Liu L, Panetta RL, Travis LD, Yang P, Zakharova NT (2016) First-principles modeling of electro-magnetic scattering by discrete and discretely heterogeneous random media. Phys Rep 632:1–75

Newton R (1969) Scattering theory of waves and particles, M, MIR

Ovchinnikov GI, Tatarsky VI (1972) To the problem on relation between coherence theory and radiative transfer equation (in Russian). Izv Vyschikch Uch Zav Radiofizika 15(9):1419–1421

Ovchinnikov GI (1974) Radiative transfer equation in visible wave band for turbulent atmosphere containing aerosol (in Russian). Izv AN SSSR, Fizika atm. i okeana 10(1):88–91

Salpeter VE, Bether H (1951) A relativistic equation for bound-state problems. Phys Rev 84:1232

Sazonov VN (1969) The generation and transfer of polarized synchrotron emission (in Russian). Astron Zhurnal 46(3):502–511

Sazonov VN (1974) Radiative transfer in anisotropic medium under the conditions of local thermodynamical equilibrium (in Russian). Astrofizika 10(3):405–416

Silant'ev NA (1971) Dissertation, Leningrad (in Russian)

Tsang L, Ishimaru A (1992) Radiative wave equations for vector electromagnetic propagation in demse nontenuous media. J Electromag Waves Appl 1:809–813

Tsang L, Kong JA (1992b) Scattering of electromagnetic waves from a dense medium consisting of correlated Mie scatterers with size distribution and applications to dry snow. J Electromag Waves Appl 6(3):265–286

Waterman PC, Truell R (1961) Multiple scattering of waves. J Math Phys 2(4):512–537

Watson KM (1969) Multiple scattering of electromagnetic waves in an underdense plasma. J Math Phys 2(4):688–702

Zheleznyakov VV, Suvorov EV, Shaposhnikov VE (1974) Transfer of polarized radiation in a magnetoactive plasma (in Russian). Astron Zhurnal 51(2):243–251

Aerosol Layer Height over Water via Oxygen A-Band Observations from Space: A Tutorial

Anthony B. Davis and Olga V. Kalashnikova

1 Introduction

Aerosols are a highly problematic area in climate science for several reasons. On the one hand, they are at least partially anthropogenic, originating from industrial facilities spewing pollution as well as agricultural activity, seasonal biomass burning, land-use change, and even wood-stove cooking in densely populated regions. On the other hand, aerosols interact in very poorly understood ways with clouds and hence, indirectly, the climate system as a whole (Boucher and Randall 2014). It is widely acknowledged that remote sensing from space will be instrumental in advancing this scientific challenge, and the National Academies have recently formulated specific ways in which NASA can to bring its satellite missions to bear on this problem in the course of the next decade (National Academies of Sciences, Engineering, and Medicine 2018).

Both active (lidar) and passive approaches will be implemented in both large (multi-sensor) and small spacecraft. One important observable is the vertical profile of aerosols in the atmosphere. This measurement is normally assigned to lidar systems because their pulsed transmitters and time-sensitive receivers can easily locate the source of the scattered laser light, assuming a single-scattering has occurred. This is however an expensive, massive, and energy-hungry technology that delivers detailed aerosol profiles down to either the surface or the highest optically thick cloud layer, along a narrow sub-satellite "curtain" (Winker et al. 2010). Over time, the curtain will sweep across all the important regions from a climate perspective. We note that there is also an acute need to determine the height of aerosol layers for the purposes of ocean color remote sensing (Duforêt et al. 2007), especially when they are absorbing and optically thin since aerosol optical thicknesses (AOTs) in excess of 0.3 are screened out.

A. B. Davis · O. V. Kalashnikova (✉)
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
e-mail: olga.v.kalashnikova@jpl.nasa.gov

There is an alternative to lidar sensing for aerosol (and cloud) profiling that is entirely passive. For starts, there is passive sensitivity to aerosol height in the near-UV via changes in the contribution of Rayleigh scattering, but multi-angle polarimetric observations are required to unravel the confounding factors of absorption and altitude(Kalashnikova et al. 2011a; Wu et al. 2016). In this study, however, we will use the abundant sunlight available in the visible/near-IR (VNIR) spectral region containing oxygen A-band (759–772 nm). This spectroscopic capability has been investigated extensively in the literature (Gabella et al. 1999; Min and Harrison 2004; Corradini and Cervino 2006; Pelletier et al. 2008; Dubuisson et al. 2009; Kokhanovsky and Rozanov 2010; Sanghavi et al. 2012; Frankenberg et al. 2012; Xu et al. 2017). Above all, we have learned that aerosol profile information content and retrieval accuracy increases with spectral resolution: as many as $\sim$2000 (Frankenberg et al. 2012) and as few as two (Xu et al. 2017) channels have been used. The later minimalistic case of just two channels, hence just one differential optical absorption spectroscopy (DOAS) ratio, is investigated here for the first time, as far as we know, using multiple view angles for aerosols.

Note that, even at high spectral resolving power, the vertical resolution of lidar is not rivaled but the passive A-band techniques apply to imaging systems. The exquisite vertical resolution achieved by lidar in its narrow linear transect through the atmosphere is therefore traded for a coarser aerosol profiling in every pixel across a potentially wide swath of imagery.

In other words, in optically thin cases (say, AOT is 0.3 at most), a lidar system can literally see through the whole layer and the so-called lidar equation (based physically on a single back-scattering) can be adopted as the forward model for the lidar return. In optically thick ones, say, AOT in excess of 3 (Vaughan et al. 2009, among others), it can not,[1] so the O_2 A-band comes to the rescue but multiple scattering needs to be accounted for in the forward model. Active lidar and passive A-band are therefore a natural synergy of observation methods. However, lidar technology is expensive and, being limited to the sub-satellite transect, the horizontal spatial sampling is inherently poor. O_2 A-band is the exact opposite on both issues, so it is important in our view to enable the passive A-band technique to embrace the whole range of optical depths, no matter if it is for aerosols or clouds, even if (as we show further on) only one piece of information is available at the small end of the AOT spectrum due to instrumental limitations.

Before narrowing our focus onto di-oxygen DOAS methods and optically thin scenarios, we note that there are other remote sensing methods for inferring aerosol layer height. For starts, the multi-angle capability of MISR (Diner et al. 1998) enables stereoscopic retrievals of heights and motion vectors for clouds (CVM) (Moroney et al. 2002), and aerosol plumes as well. While cloud height are reported operationally as part of MISR "CVM" product, the plume height data for aerosols from MISR is derived manually with the MISR Interactive eXplorer (MINX) tool (Nelson et al.

---

[1]Davis et al. (2009), and authors cited therein) take the opposite perspective and argue that, with *new* lidar equations that account fully for multiply scattered laser light, lidar techniques can be extended to optically thick clouds and, for that matter, aerosol plumes.

2008, 2013). By incorporating additional information on the direction of apparent plume motion from a trained user, the MINX visualization and analysis software enables precise retrievals of aerosol plume heights and instantaneous winds at a horizontal resolution of 1.1 km. MINX plume height and motion retrievals provided unique and valuable information on aerosol injection heights from volcanic eruptions (Flower and Kahn 2018), wild fires (Val Martin et al. 2010), and dust source activations (Kalashnikova et al. 2011b).

Another remote sensing take on aerosol layer height from multi-angle imaging is the emerging class of aerosol plume tomography techniques. When the goal is to reconstruct the 3D structure of an an aerosol plume, part of that is its top and bas altitude. Aides et al. (2013) demonstrate an efficient physics-based reconstruction of synthetic aerosol distributions. Like here, but in 3D, they adopt a single-scattering radiative transfer (RT) model to solve the very high dimensional inverse problem on a 3D grid of unknown opacity values, and they quantify at least forward modeling error when plume opacity is increased to the point where multiple scattering becomes significant. In contrast, Garay et al. (2016) demonstrate the reconstruction of a real-world smoke plume sufficiently far from the wild fire source that it is optically thin. These last authors used a combination of MISR's geolocated radiances for nadir and all oblique views ("Level 1b" data) and retrieved ("Level 2") aerosol properties, and validated their transect reconstruction by comparison with aerosol profiling from a nearby lidar station. In fact, due to the strong need for "regularization" to make the large inverse problem tractable, Garay et al.'s most robust retrieval was indeed the tenuous smoke plume's smooth top altitude.

The remainder of the paper is structured as follows. In Sect. 2, we describe our forward model for DOAS radiance ratios based on the O_2 A-band as measured from a space-based sensor. In Sect. 3, a simplified version of this RT problem is described that is tailored specifically for information content (IC) analysis, and not for real-world data analysis; it is then solved analytically. We are then in a position to derive in closed-form the associated Jacobian matrix elements for DOAS ratios for the simplified forward RT model. In Appendix "IC Analysis in Optimal Estimation Theory", we summarize signal IC analysis in the general framework of optimal estimation (OE), which makes extensive use of Jacobians. In Sect. 4, we reckon with the instrumental idiosyncrasies for the Ocean Color Imager (OCI) developed at NASA/GSFC for the core spectrometry on the PACE (Plankton, Aerosol, Cloud, and ocean Ecosystem) mission (NASA 2017) and/or a multi-angle sensor (MAS) since such sensors are being considered for the same platform.[2] In Sect. 5, we perform quantitative assessments of IC for mono-angle spectrometry, for multi-angle sampling of a single DOAS ratio, and for the fusion of both modalities. Finally, in Sect. 6, we summarize and qualify our findings, and outline future work.

Remark The present paper is written as a tutorial that documents key aspects of a demonstration of feasibility for proposed retrievals of a single-layer (hence two-parameter) aerosol profile using the O_2 A-band. In particular, it transpires that,

---

[2]These "contributed" sensors for the PACE mission are in fact imaging polarimeters, but we will only use the intensity (i.e., 1st component of the Stokes vector) herein.

between the first physics-based modeling of the differential optical absorption spectroscopy (DOAS) signals using simplified one-dimensional (1D) RT and the following OE-based analysis of IC about the parameterized aerosol profile, an unexpectedly large sample of the math encountered in a typical university curriculum is exploited. We will list the topics used in the Summary section, and discuss the possibility of using this study as a showcase of "applicable" math in a single remote sensing research project. In a related publication (Davis et al. 2018), we will use a computational 1D RT model to quantify the modeling error associated with the simplifying assumptions made about the optical properties of the atmosphere/surface system in order to reduce herein the model to a closed-form expression, as is required for the tutorial exercise. We will also highlight the application of the simplified DOAS signal model to the problem of atmospheric correction in ocean color studies.

2 1D RT Modeling for DOAS Ratios

Let $I_\lambda(z, \mathbf{\Omega})$ be radiance at wavelength λ in direction $\mathbf{\Omega}$ at level z in a plane-parallel optical medium $M = \{\mathbf{x} = (x, y, z)^T \in \mathbb{R}^3;\ 0 \leq z \leq H\}$, where we can take $H = \infty$, as needed. M is characterized optically by height-dependent scattering, absorption and extinction coefficients, denoted respectively $\sigma_s(z)$, $\sigma_{a\lambda}(z)$ and $\sigma_\lambda(z)$, in m^{-1}, where

$$\sigma_\lambda(z) = \sigma_s(z) + \sigma_{a,\lambda}(z). \tag{1}$$

We also need to specify the phase function $P(z, \mathbf{\Omega} \cdot \mathbf{\Omega}')$ in sr^{-1} (hence normalized to unity) for scattering from $\mathbf{\Omega}$ into $\mathbf{\Omega}'$, or vice versa. We allow for extinction to vary with λ, but not the scattering properties $\sigma_s(z)$ and $P(z, \mathbf{\Omega} \cdot \mathbf{\Omega}')$ because, in the present study at least, we focus on a narrow spectral range where only molecular absorption cross-section varies significantly.

Furthermore, we identify two sources of light absorption:

$$\sigma_{a,\lambda}(z) = \sigma_a^{(p)}(z) + k_\lambda(z), \tag{2}$$

where $\sigma_a^{(p)}(z)$ is the particulate (aerosol) absorption coefficient and $k_\lambda(z)$ is the interstitial gaseous absorption coefficient.

Figure 1 shows our detailed representation of the O_2 A-band region at a resolving power $\lambda/\Delta\lambda = 15300$ using total-column absorption optical depth,

$$\tau_{O_2}(\lambda) = \int_0^H k_\lambda(z)\, dz. \tag{3}$$
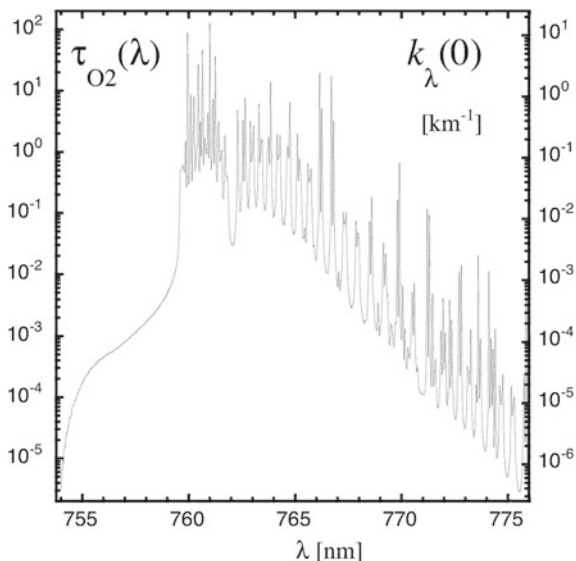
Fig. 1 τ_{O_2} in (3) versus λ at resolution $\Delta\lambda = 0.05$ nm. At this resolution, maximum O_2 optical depth is about 125. Assuming, for simplicity, an approximately exponential profile for of absorption coefficient in altitude, i.e., $k_\lambda(z) \approx k_\lambda(0)\exp(-z/H_{mol})$, with a pressure scale height $H_{mol} = 8$ km, we have $k_\lambda(0) \approx \tau_{O_2}(\lambda)/H_{mol}$, which is tracked on the right-hand axis. These data were kindly provided to us by Dr. Suniti Sanghavi, who generated them from the HITRAN line-by-line database (Rothman 2010). This resolution yields a maximum ground-level extinction of $k_\lambda(0) = 15.6$ km^{-1}, hence a mean-free-path (also e-folding distance) $\langle\ell\rangle = 1/k_\lambda(0) \approx 64$ m for direct transmission based on Beer's law: $T_\lambda(\ell) = e^{-k_\lambda \ell}$. Indeed, viewed physically as a cumulative probability of surviving extinction over a distance ℓ, it yields $\langle\ell\rangle_\lambda = \int_0^\infty \ell\,|dT_\lambda(\ell)/d\ell|\,d\ell = 1/k_\lambda$

Radiance $I_\lambda(z, \boldsymbol{\Omega})$, which has units of W/m^2/sr/nm, obeys the 1D RT equation (RTE)[3]

$$\mu \frac{dI_\lambda}{dz} + \sigma_\lambda(z)I_\lambda = S_\lambda(z, \boldsymbol{\Omega}) + q_\lambda(z, \boldsymbol{\Omega}), \tag{4}$$

where $\mu = \Omega_z$. The above integro-differential 1D RTE balances sinks (left-hand side), namely, directional change and extinction, and sources (right-hand side) at position $(z, \boldsymbol{\Omega})$ in the composite space M×unit sphere (denoted hereafter as usual by "4π"). Accordingly, $q_\lambda(z, \boldsymbol{\Omega})$ is the known source *term*, in W/m^3/sr/nm, describing the injection of radiant energy at $(z, \boldsymbol{\Omega})$, and

$$S_\lambda(z, \boldsymbol{\Omega}) = \sigma_s(z) \int_{4\pi} P(z, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}')I_\lambda(z, \boldsymbol{\Omega}')d\boldsymbol{\Omega}' \tag{5}$$

---

[3]Although long-accepted on a phenomenological basis (Chandrasekhar 1950), the RTE in (4)–(5) was derived rigorously from EM wave theory (i.e., Maxwell's equations) and statistical optics only in 2002 by Mishchenko Mishchenko (2002).

is the generally unknown source *function* describing in-scattering from any direction into $\mathbf{\Omega}'$ at level z.

It is convenient in numerical transport theory to change the vertical coordinate from z to

$$\tau_z(\lambda) = \int\limits_z^H \sigma_\lambda(z')\,\mathrm{d}z', \tag{6}$$

which varies over a finite interval $[0,\tau_0(\lambda)]$. We can thus define the inverse $z_\tau(\lambda)$ of this function, which relates $[0,\tau_0(\lambda)]$ to $[0,H]$ for a given λ, potentially for $H = \infty$. We need to know this correspondence since optical properties are generally prescribed in z. For ease of reading, we drop in the remainder the subscript z and the argument λ in (6). We then have to solve the coupled equations

$$\mu\,\frac{\mathrm{d}I_\lambda}{\mathrm{d}\tau} + I_\lambda(\tau, \mathbf{\Omega}) = S_\lambda(\tau, \mathbf{\Omega}) + q_\lambda(\tau, \mathbf{\Omega}), \tag{7}$$

$$S_\lambda(\tau, \mathbf{\Omega}) = \omega_\lambda(\tau) \int\limits_{4\pi} P(\tau, \mathbf{\Omega} \cdot \mathbf{\Omega}')I_\lambda(\tau, \mathbf{\Omega}')\mathrm{d}\mathbf{\Omega}' \tag{8}$$

for $I_\lambda(\tau, \mathbf{\Omega})$, where

$$\omega_\lambda(\tau) = \frac{\sigma_\mathrm{s}(z_\tau)}{\sigma_\lambda(z_\tau)} \tag{9}$$

is the layer's SSA. For completion, we need to specify boundary conditions (BCs) for (7), that is, $I_\lambda(0, \mathbf{\Omega})$ for $\mu > 0$ and $I_\lambda(\tau_0, \mathbf{\Omega})$ for $\mu < 0$.

Consider now the so-called "formal solution" of (7), which is best formulated separately for downwelling ($\mu > 0$) and upwelling ($\mu < 0$) radiances:

$$I_\lambda(\tau, \mathbf{\Omega}) = \begin{cases} \int_0^\tau [S_\lambda + q_\lambda]\left(\tau', \mathbf{\Omega}\right) \mathrm{e}^{-\frac{\tau-\tau'}{\mu}}\mathrm{d}\tau'/\mu + I_\lambda(0, \mathbf{\Omega})\mathrm{e}^{-\frac{\tau}{\mu}}, & \text{for } \mu > 0 \\ \int_\tau^{\tau_0} [S_\lambda + q_\lambda]\left(\tau', \mathbf{\Omega}\right) \mathrm{e}^{-\frac{\tau'-\tau}{|\mu|}}\mathrm{d}\tau'/|\mu| + I_\lambda(\tau_0, \mathbf{\Omega})\mathrm{e}^{-\frac{\tau_0-\tau}{|\mu|}}, & \text{for } \mu < 0 \end{cases}. \tag{10}$$

Substitution of (8) into (10) yields an integral 1D RTE that can be solved numerically by iteration. In operator short-hand, we have (Davis and Knyazikhin 2005)

$$I_\lambda = \mathcal{K}_\lambda * I_\lambda + Q_\lambda, \tag{11}$$

where Q_λ is a generic source term, $\mathcal{K}_\lambda$ is the kernel of the integral equation, and "$*$" means a convolution product. In the case of 1D RT, this convolution combines both the angular integration in (8) and the "up-wind" spatial integration in (10) where, logically, contributions to radiance somewhere, in a certain direction $\mathbf{\Omega}$, can only come from sources (including in-scattering) distributed along $-\mathbf{\Omega}$.

This leads to a straightforward numerical solution by Neumann series expansion (Evans and Marshak 2005), a.k.a. successive orders of scattering:

$$I_\lambda(\tau, \boldsymbol{\Omega}) = (\mathcal{E} - \mathcal{K})^{-1} * Q_\lambda(\tau, \boldsymbol{\Omega})$$

$$= \sum_{n=0}^{\infty} I_\lambda^{(n)}(\tau, \boldsymbol{\Omega}) = \sum_{n=0}^{\infty} \mathcal{K}_\lambda^n(\tau, \boldsymbol{\Omega}; \tau', \boldsymbol{\Omega}') * I_\lambda^{(0)}(\tau', \boldsymbol{\Omega}'), \qquad (12)$$

where $\mathcal{E}$ is the identity operator, and we have introduced the *directly* transmitted radiance field $I_\lambda^{(0)}(\tau, \boldsymbol{\Omega}) = Q_\lambda(\tau, \boldsymbol{\Omega})$ along with the iteration rule $I_\lambda^{(n+1)}(\tau, \boldsymbol{\Omega}) = \mathcal{K}_\lambda(\tau, \boldsymbol{\Omega}; \tau', \boldsymbol{\Omega}') * I_\lambda^{(n)}(\tau', \boldsymbol{\Omega}')$.

For solar illumination, we take $q_\lambda(\tau, \boldsymbol{\Omega}) \equiv 0$ in (7), hence in (10), and set the upper boundary condition (BC) in (10, top option) as:

$$I_\lambda(0, \boldsymbol{\Omega}) = F_{0\lambda}\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}_0), \text{ for } \mu > 0, \qquad (13)$$

where F_0 denotes the incoming spectral solar irradiance, in W/m^2/nm and $\boldsymbol{\Omega}_0$ is the incoming solar direction with $\Omega_{0z} = \mu_0 = \cos\theta_0 > 0$, where θ_0 is SZA; we can set $\Omega_{0x} = \sqrt{1 - \mu_0^2}$ and $\Omega_{0y} = 0$ (i.e., $\phi_0 = 0$) without loss of generality. The associated source term for the integral equation in (11), and to be used in its iterative solution (12), is

$$Q_\lambda(\tau, \boldsymbol{\Omega}) = I_\lambda^{(0)}(\tau, \boldsymbol{\Omega}) = \begin{cases} F_{0\lambda}e^{-\frac{\tau}{\mu_0}}, \text{ for } \mu = \mu_0 \text{ and } \phi = \phi_0 \\ 0, \qquad\qquad \text{otherwise} \end{cases}, \qquad (14)$$

which results from the 2nd term in (10, upper option) using the upper BC expressed in (13).

A standard lower BC for (10, bottom option) describes bi-directional reflectivity:

$$I_\lambda(\tau_0, \boldsymbol{\Omega}) = \int_{\mu'>0} \rho_\lambda(\boldsymbol{\Omega}', \boldsymbol{\Omega})I_\lambda(\tau_0, \boldsymbol{\Omega}')d\boldsymbol{\Omega}', \text{ for } \mu < 0, \qquad (15)$$

where $\rho_\lambda(\boldsymbol{\Omega}', \boldsymbol{\Omega})$ is the bi-directional reflection distribution function (BRDF) of the underlying surface. We note that the surface BDRF $\rho_\lambda(\boldsymbol{\Omega}', \boldsymbol{\Omega})$ in the right-hand side is the target in atmospheric correction for landmass remote sensing.

3 Simplified RT Modeling for Satellite DOAS Signals

3.1 Simplifying Assumptions

First, we will assume a black, purely absorbing surface. In reality, water is a highly specular reflector, even ruffled by wind, but we will not consider sunglint directions. Barring high surface wind (hence white caps and broadened sunglint) and turbidity
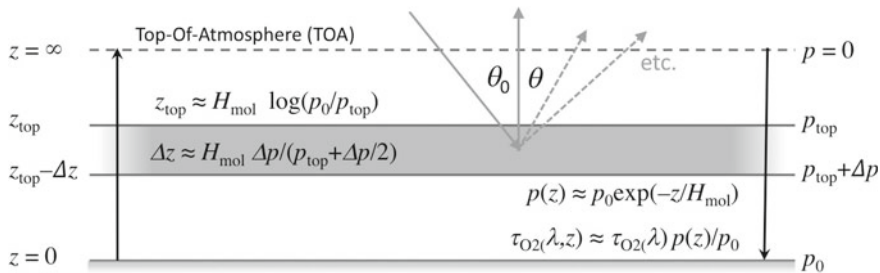
Fig. 2 Schematic of 3-layer optical medium used in this study, along with various definitions and approximations used in the study. Note that, in the limits of no scattering outside of the aerosol layer and no surface reflection, only top and middle layers matter

("case-2" water), the ocean's reflectivity is very low indeed, especially in the red-to-NIR transition region where the O_2 A-band resides.

Secondly, we will focus on the dependence of O_2's absorption cross-section C_{O_2} on wavelength λ, and neglect its temperature and pressure dependencies. That leaves the gaseous absorption coefficient, $k_\lambda(z)$ expressed as $C_{O_2}(\lambda) \times$ molecular density, which is proportional to atmospheric pressure $p(z)$. As needed, we will further assume that the pressure profile is exponential, with a scale-height $H_{mol} \approx 8$ km: $p(z) \approx p_0 \exp(-z/H_{mol})$, where p_0 sea-level pressure that, for simplicity, we will assume is standard (1013.25 hPa).

Our third simplification is to divide the atmosphere into two or three distinct layers: one above the aerosols (with absorbing gas only), a mixed layer of uniform aerosols and absorbing gas and, as needed, one below the aerosols (again absorbing gas only). See schematic in Fig. 2. This is to say that we are concentrating the scattering by aerosol particles in one well-defined layer. In fact, the two key remote sensing unknowns for us are p_{top}, the pressure at the top of the aerosol layer, and Δp, pressure thickness of the layer. As needed, these key parameters can be translated into layer-top altitude, $z_{top} \approx H_{mol} \log(p_0/p_{top})$, and geometric layer thickness, $\Delta z \approx H_{mol} \times \Delta p/(p_{top} + \Delta p/2)$, assuming $\Delta p \ll p_{top}$ and invoking the mean-value theorem for improved accuracy.

Our fourth simplification is that aerosol optical thickness (AOT) at the A-band, τ_p, remains $\lesssim 0.1$. This low-AOT assumption will enable us to invoke the single-scattering approximation for the RT problem, which leads to relatively simple closed-form expressions. This constraint roughly translates to an AOT (usually at 550 nm) in the middle of the visible (VIS) spectrum that is $\lesssim 0.15$ to 0.2, assuming the Angstrom exponent ranges between 1 (coarse particulates) and 2 (finer particles). These mid-VIS AOTs are typical of conditions where the atmospheric correction of PACE radiances at the top-of-atmosphere (TOA) should perform well enough to deliver the desired water-leaving radiance.

At A-band wavelengths, Rayleigh optical thickness of the total atmospheric column is only ≈ 0.025, i.e., $1/4$ or so of the nominal AOT budget of ≈ 0.1. Rayleigh scattering is therefore neglected as a fifth simplification. We are confident that this is

a tolerable modeling error for our present purposes of *rough* IC analysis, that is, provide yes/no answers to the question of retrievability of p_{top} and/or Δp. However, this simplification would introduce a systematic bias in an actual retrieval algorithm that is unnecessary, even in the analytical single-scattering approximation. Note that, in the absence of Rayleigh scattering and of surface reflection, the optional atmospheric layer below the aerosols is immaterial until one of these hypothesis is relaxed.

Finally, a sixth simplification is introduced further on to deal with spectral integration by the space-based instruments that we will investigate.

3.2   Derivation of an Analytical Expression for DOAS Ratios

In the above, we make five simplifying assumptions that enable a closed-form analytical solution of the DOAS RT problem in a plane-parallel atmosphere. Briefly, they are:

1. surface is black, i.e., BRDF $\rho_\lambda(\mathbf{\Omega}', \mathbf{\Omega})$ in (15) vanishes identically;
2. detailed variations of O_2 absorption cross-section with temperature and pressure are negligible, which enables us to write the gaseous absorption optical depth from TOA to pressure level $p \leq p_0$ as

$$\tau_{\text{g}}(\lambda, p) \approx \tau_{O_2}(\lambda) \times p/p_0, \tag{16}$$

   with $\tau_{O_2}(\lambda)$ from (3);
3. atmosphere has just 2 or 3 layers (from TOA down to surface, absorbing gas, mixed aerosol and gas, more gas if the aerosols are primarily lofted above the boundary layer);
4. aerosol layer is optically thin at the A-band, i.e., $\tau_{\text{p}} = \sigma_{\text{p}}\Delta z$, remains $\lesssim 0.1$, which translates to AOT $\lesssim 0.15$ to 0.2 in the mid-VIS depending on Angstrom exponent;
5. Rayleigh scattering is negligible, as the associated optical depth at the A-band is only ≈ 0.025.

We thus neglect hereafter all orders of scattering $n \geq 2$ in the expansion of radiance in (12). Directly transmitted (a.k.a. "uncollided") radiance is already given in (14). The singly scattered radiance in any upward direction is obtained by applying the integral kernel K_λ once to it. In other words, we substitute (14), i.e., $F_{0\lambda}e^{-\frac{\tau}{\mu_0}}\delta(\mathbf{\Omega} - \mathbf{\Omega}_0)$, into the definition (8), and then use the result in (10, lower option), with both $q_\lambda(\tau', \mathbf{\Omega})$ and $I_\lambda(\tau_0, \mathbf{\Omega})$ set to 0. The outcome is

$$I_\lambda^{(1)}(\tau, \mathbf{\Omega}) = F_{0\lambda} \int\limits_{\tau}^{\tau_0(\lambda)} \omega_\lambda(\tau')P(\tau', \mathbf{\Omega}_0 \cdot \mathbf{\Omega})e^{-\left(\frac{1}{\mu_0} + \frac{1}{|\mu|}\right)\tau'}d\tau'/|\mu|. \tag{17}$$

In the remainder, we will focus on upwelling radiance escaping to space (i.e., $\tau = 0$ in the above) in direction $\mathbf{\Omega}(\mu, \phi)$ where $\mu < 0$ and ϕ is the azimuthal angle relative

to the principal plane (containing the vertical axis and $\mathbf{\Omega}_0$); in particular, we will be using $|\mu| = \cos\theta$, where θ is VZA.

We now invoke the assumed 2- or 3-layered structure in Fig. 2, reduced effectively to just the top two layers by neglecting surface reflection and Rayleigh scattering:

1. A purely absorbing ($\omega_\lambda(\tau) \equiv 0$) gaseous layer between TOA at $p = 0$ (also $z = H = \infty$, $\tau = 0$) and the top of the scattering layer at $p = p_{\text{top}}$, hence $\tau_{\text{top}}(\lambda) = \tau_{O_2}(\lambda) \times p_{\text{top}}/p_0$ from (16), effectively the lower bound of the integral in (17). Its contribution to (17) is therefore a multiplicative factor of

$$T_{\text{above}}^{(2\text{-way})}(\tau_{O_2}(\lambda); \mu_0, |\mu|; p_{\text{top}}) = e^{-\left(\frac{1}{\mu_0} + \frac{1}{|\mu|}\right)\tau_{\text{top}}(\lambda)} = e^{-\left(\frac{1}{\mu_0} + \frac{1}{|\mu|}\right)\tau_{O_2}(\lambda)\frac{p_{\text{top}}}{p_0}}. \quad (18)$$

2. A mixed aerosol and gas layer where we have a uniform phase function value, hence another multiplicative factor $P(\mathbf{\Omega}_0 \cdot \mathbf{\Omega})$, extending between $\tau_{\text{top}}(\lambda)$ and $\tau_{\text{top}}(\lambda) + \tau_{\text{tot}}(\lambda)$, where $\tau_{\text{tot}}(\lambda) = \tau_p + \tau_{O_2}(\lambda)\Delta p/p_0$. We assume that its pressure thickness Δp is small enough compared to p_{top} and p_0 that we can take the gaseous absorption coefficient k_λ as constant in the layer. If so, then the SSA term in (17) is uniform inside the layer, and given by

$$\omega_{\text{mixed}}(\tau_{O_2}(\lambda); \Delta p; \omega_p, \tau_p) = \frac{\omega_p \tau_p}{\tau_{\text{tot}}(\lambda)} = \frac{\omega_p \tau_p}{\tau_p + \tau_{O_2}(\lambda)\Delta p/p_0}, \quad (19)$$

   where ω_p is the SSA for the aerosol particles.

With the change of variables $\tau' = \tau_{\text{top}}(\lambda) + \tau''$, Eq. (17) at $\tau = 0$ becomes:

$$I_\lambda^{(1)}(0, \mathbf{\Omega}) = F_{0\lambda}\, T_{\text{above}}^{(2\text{-way})}(\tau_{O_2}(\lambda); \mu_0, |\mu|; p_{\text{top}}) \quad (20)$$

$$\times\, \omega_{\text{mixed}}(\tau_{O_2}(\lambda); \Delta p; \omega_p, \tau_p)\, P(\mathbf{\Omega}_0 \cdot \mathbf{\Omega}) \int\limits_0^{\tau_{\text{tot}}(\lambda)} e^{-\left(\frac{1}{\mu_0} + \frac{1}{|\mu|}\right)\tau''}\mathrm{d}\tau''/|\mu|,$$

where $\tau_{\text{tot}}(\lambda)$ is defined as the denominator in (19), and the integral can be evaluated analytically.

Overall, the scattering layer contributes the 2nd line in (20), a multiplicative factor that is best written as $(\mu_0/\pi) R(\mathbf{\Omega}_0, \mathbf{\Omega}; \omega_p, \tau_p; \tau_{O_2}(\lambda); \Delta p)$, where $R(\cdots)$ is the layer's bi-directional reflection factor (BRF). Collecting expressions from (18)–(20), the observed signal in "BRF" form is

$$\frac{\pi I_\lambda^{(1)}}{\mu_0 F_{0\lambda}}(0, \mathbf{\Omega}) = T_{\text{above}}^{(2\text{-way})}(\tau_{O_2}(\lambda); \mu_0, |\mu|; p_{\text{top}})\, R(\tau_{O_2}(\lambda); \mathbf{\Omega}_0, \mathbf{\Omega}; \Delta p; \omega_p, \tau_p) \quad (21)$$

where

$$R(\tau_{O_2}(\lambda); \boldsymbol{\Omega}_0, \boldsymbol{\Omega}; \Delta p; \omega_p, \tau_p) = \omega_{\text{mixed}}(\tau_{O_2}(\lambda); \Delta p; \omega_p, \tau_p) \left[4\pi P(\boldsymbol{\Omega}_0 \cdot \boldsymbol{\Omega})\right]$$
$$\times \frac{1 - e^{-\left(\frac{1}{\mu_0} + \frac{1}{|\mu|}\right)\left(\tau_p + \tau_{O_2}(\lambda)\frac{\Delta p}{p_0}\right)}}{4(\mu_0 + |\mu|)}. \tag{22}$$

The term in square brackets in the 1st line is the popular non-dimensional way of writing the phase function, which gives the more familiar functional form used in the 2nd line.

In the last five numbered equations, we have isolated with semi-colons variables from: (i) O_2, i.e., spectroscopy; (ii) solar and viewing geometry; (iii) aerosol profile, which is the A-band remote sensing unknown; and (iv) aerosol properties that need to be determined from other observations.

We can now form the desired DOAS ratio by dividing (21)–(22) by its "continuum" value at $\tau_{O_2}(\lambda) = 0$. Letting $M(\mu_0, |\mu|) = \frac{1}{\mu_0} + \frac{1}{|\mu|}$ denote airmass, this works out to be

$$\frac{I_\lambda^{(1)}}{I_{\text{cont}}^{(1)}}(\tau_{O_2}(\lambda); M; p_{\text{top}}, \Delta p; \tau_p) = e^{-M\tau_{O_2}(\lambda)\frac{p_{\text{top}}}{p_0}} \tag{23}$$

$$\times \frac{1 - e^{-M\left(\tau_p + \tau_{O_2}(\lambda)\frac{\Delta p}{p_0}\right)}}{1 - e^{-M\tau_p}} \times \frac{\tau_p}{\tau_p + \tau_{O_2}(\lambda)\frac{\Delta p}{p_0}},$$

which is independent of the aerosol's SSA ω_p and phase function $4\pi P(\boldsymbol{\Omega}_0 \cdot \boldsymbol{\Omega})$. We note that ω_p must nonetheless be non-zero, and the closer to unity the better (to boost signal). The 1st line in (23) comes from the upper layer of absorbing gas, and the 2nd line in (23) is from the ratio of aerosol layer BRFs with the absorbing gas mixed in divided by the same expression without the absorbing gas.

3.3 Simplified 1D RT Forward Model in Final Form

To the extent that the end justifies the means, our five simplifying assumptions have lead to a relatively simple expression in (23) for the DOAS signals of interest here that is simple enough to analyze by inspection. We thus gain intuitive physical insights into the passive remote sensing of aerosol height using O_2 A-band measurements from space, on top of performing the formal—and somewhat abstract—statistics-based IC analysis further on.

The monochromatic DOAS ratio obtained in (23) can be redefined as our forward model for remote sensing signals:

$$F_\lambda(M; \mathbf{x}; \mathbf{b}) = F_\lambda(M; p_{\text{top}}, \Delta p; \tau_{\text{p}}) \tag{24}$$

$$= e^{-M\tau_{\text{O}_2}\frac{p_{\text{top}}}{p_0}} \left(\frac{1 - e^{-M\left(\tau_{\text{p}} + \tau_{\text{O}_2}\frac{\Delta p}{p_0}\right)}}{1 - e^{-M\tau_{\text{p}}}} \right) \left(\frac{\tau_{\text{p}}}{\tau_{\text{p}} + \tau_{\text{O}_2}\frac{\Delta p}{p_0}} \right),$$

where we use the airmass factor

$$M(\mu_0, |\mu|) = \frac{1}{\mu_0} + \frac{1}{|\mu|} = \frac{1}{\cos\theta_0} + \frac{1}{|\cos\theta|}, \tag{25}$$

with θ_0 and θ denoting respectively the solar and viewing zenith angles (SZA and VZA). All the dependence on λ is via τ_{O_2}. The OCI spectrometer modulates the DOAS ratio via $\tau_{\text{O}_2}(\lambda)$, while the MAS does it by changing $\mu(\theta)$ in M. By sampling either λ or M, we build up the n-dimensional vector-valued $\mathbf{F}_\lambda(M; p_{\text{top}}, \Delta p; \tau_{\text{p}})$ used further on. The m-dimensional "state space" is spanned by $\mathbf{x} = (p_{\text{top}}, \Delta p)$ in this study, hence $m = 2$, and the non-retrieved parameter space contains only $\mathbf{b} = \tau_{\text{p}}$.

A first-order series expansion of the approximate DOAS ratio in (24) in both $\tau_{\text{O}_2}(\lambda)$ and τ_{p}, which contains the first cross-term in $\tau_{\text{p}}\tau_{\text{O}_2}(\lambda)$, yields[4]

$$F_\lambda(M; p_{\text{top}}, \Delta p; \tau_{\text{p}}) \approx 1 - \frac{p_{\text{top}} + \left(1 - M\tau_{\text{p}}/6\right)\Delta p/2}{p_0} M\tau_{\text{O}_2}(\lambda) + \cdots \tag{26}$$

We note that the correction in $M\tau_{\text{p}}/6$ is necessarily quite small because (*i*) τ_{p} is already small, and (*ii*) we want to stay away from large airmasses; on both accounts, this safeguards the single-scattering approximation. It therefore makes physical sense that the cumulative optical path in the oxygen is $\approx M\tau_{\text{O}_2}(\lambda)(p_{\text{top}} + \Delta p/2)/p_0$, i.e., down from the TOA to the *middle* of the aerosol layer and back to space.

Figure 3 illustrates our simplified estimation of the monochromatic DOAS ratio in (24) for typical solar/viewing geometry as a function of $(p_{\text{top}}, \Delta p)$ for different values of $\tau_{\text{O}_2}(\lambda)$. Notice that $\tau_{\text{O}_2}(\lambda)$ must become quite large before lines of constant DOAS ratio start gaining any significant curvature, at which point there may not be much light left (low SNR). That difference in local slope is key to retrieving both height and thickness of the aerosol layer. This prediction is confirmed in the next Sections where the OCI spectrometer under consideration will not have access to deep enough spectral lines, for lack of resolving power.

More importantly, the slopes of the nearly straight iso-DOAS ratio lines need to change significantly from one observation to the next if we want to retrieve both p_{top} and Δp. From (26), we see that it can only happen by modulating $M(\mu_0, |\mu|)$ in multi-angle mode, and not by changing τ_{O_2}, via λ, in spectroscopy. This too will be confirmed numerically in the next Sections.

---

[4]Because of cancellations between the middle and last factors in the 2nd line of (24), the exponentials in the middle term have to be expanded to order 3 to obtain the first surviving cross-term in $\tau_{\text{p}}\tau_{\text{O}_2}(\lambda)$.
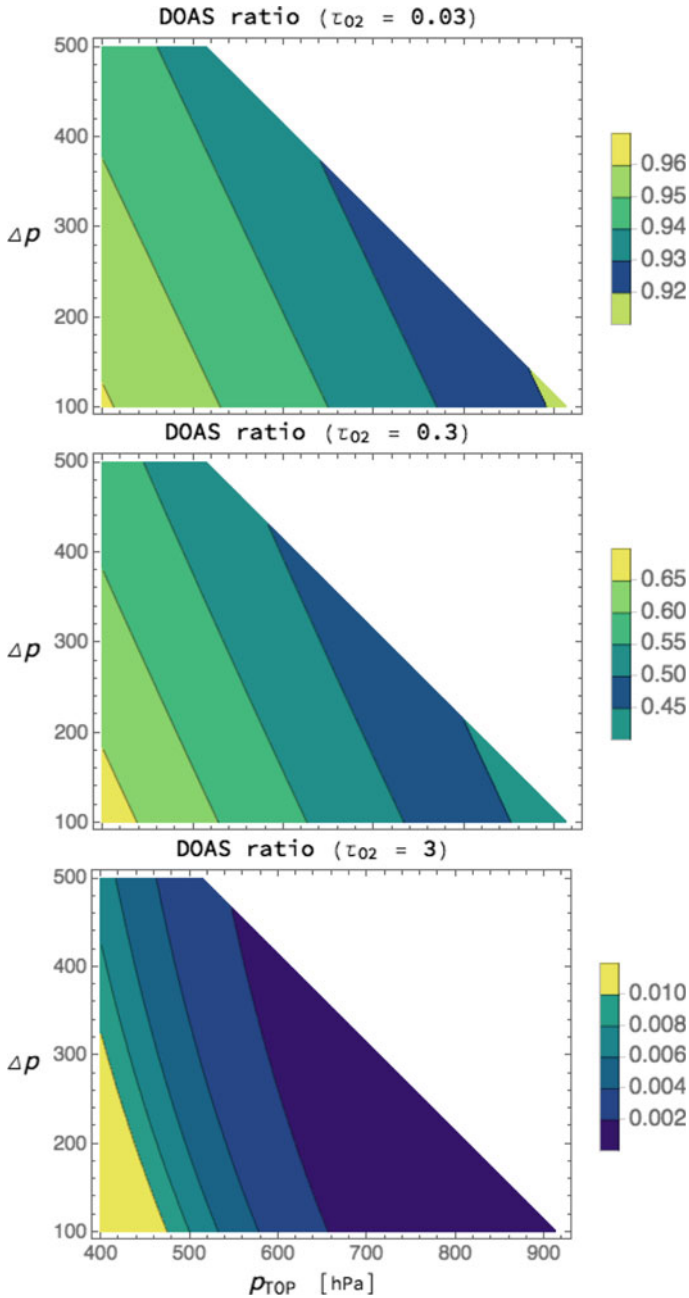
Fig. 3 DOAS ratio in (24) as a function of $(p_{\text{top}}, \Delta p)$ for $(\mu_0, |\mu|) = (1/2, 1)$, hence airmass $M = 3$ in (25), and $\tau_{\text{p}} = 0.1$. From left to right, we have $\tau_{O_2} = 0.03, 0.3, 3$. Note that the iso-ratio curves only start deviating from nearly straight lines, where $p_{\text{top}} + \Delta p/2(1 - M\tau_{\text{p}}/6) \approx$ constant, according to (26), at quite large values of τ_{O_2}

3.4 Jacobians of Simple 1D RT Model

To conduct our IC analysis further on, we will use Rodgers' OE theory (Rodgers 2000), which calls for Jacobian matrices, namely, $\partial \mathbf{F}_\lambda / \partial \mathbf{x}$ and $\partial \mathbf{F}_\lambda / \partial \mathbf{b}$. In the present case, Jacobians are more readily expressed as logarithmic partial derivatives, $\partial \ln F_\lambda / \partial \ln x_i = (x_i / F_\lambda) \times \partial F_\lambda / \partial x_i$ $(i = 1, \ldots, m)$, and similarly for the components of $\mathbf{b}$.

For the model in (24), these logarithmic Jacobians are:

$$\frac{p_{\text{top}}}{F_\lambda} \frac{\partial F_\lambda}{\partial p_{\text{top}}} = -M \tau_{O_2} \frac{p_{\text{top}}}{p_0}; \tag{27}$$

$$\frac{\Delta p}{F_\lambda} \frac{\partial F_\lambda}{\partial \Delta p} = -\left(\frac{1}{M(\tau_p + \tau_{O_2} \Delta p / p_0)} - \frac{1}{e^{M(\tau_p + \tau_{O_2} \Delta p / p_0)} - 1} \right) M \tau_{O_2} \frac{\Delta p}{p_0}; \tag{28}$$

$$\frac{\tau_p}{F_\lambda} \frac{\partial F_\lambda}{\partial \tau_p} = \left(\tau_{O_2} \frac{\Delta p}{p_0} \right.$$
$$+ e^{M \tau_p} \left[\left((e^{M \tau_p} - 1) e^{M \tau_{O_2} \Delta p / p_0} - 1 \right) \tau_{O_2} \frac{\Delta p}{p_0} \right.$$
$$\left. + \left(e^{M \tau_{O_2} \Delta p / p_0} - 1 \right) \left(\tau_p + \tau_{O_2} \frac{\Delta p}{p_0} \right) M \tau_p \right] \right)$$
$$\left/ \left[\left(e^{M \tau_p} - 1 \right) \left(e^{M(\tau_p + \tau_{O_2} \Delta p / p_0)} - 1 \right) \left(\tau_p + \tau_{O_2} \frac{\Delta p}{p_0} \right) \right]. \right. \tag{29}$$

It can be shown that the first two expressions are always negative. Interestingly, $\partial \ln F_\lambda / \partial p_{\text{top}}$ in (27) is independent of p_{top}, Δp and τ_p.

Figure 4 shows these three sensitivities as functions of the two variables sampled by different types of sensor, namely, τ_{O_2} and M. A spectrometer, like OCI/PACE, samples τ_{O_2} via wavelength λ. A multi-angle imager, like the potential MAS/PACE, samples airmass M. From left to right, we see the logarithmic derivatives of the DOAS ratio $F_\lambda(M; \cdots)$ with respect to p_{top}, Δp, and τ_p for a lofted aerosol layer between 3 and 5 Km altitude, i.e. $\{p_{\text{top}}, \Delta p\} = \{542,154\}$ hPa, with $\tau_p = 0.1$. The relative magnitudes in the three legends make clear what the easiest property to retrieve is by either observational modality, namely, p_{top}. There is much less sensitivity to Δp unless one can access large values of τ_{O_2}, meaning high spectral resolution. There is very little sensitivity to τ_p in the DOAS ratio, which is good news since it is a required parameter here, but that is retrieved otherwise, with known uncertainty. We see here that this uncertainty will not affect the DOAS-based retrievals very much.

4 Impacts of Instrument Design

We will discuss passive aerosol profiling capability using the O_2 A-band for two types of instrument.
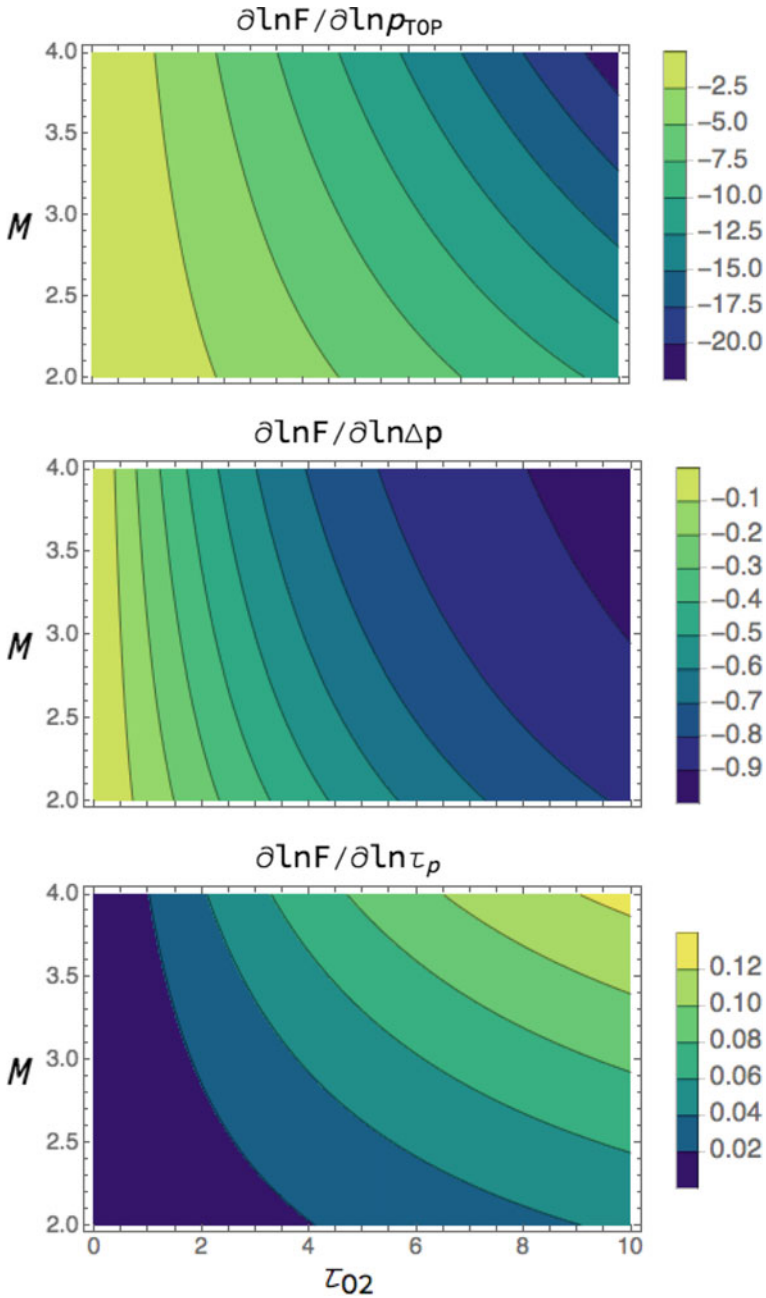
Fig. 4 Logarithmic derivatives of the forward model $F_\lambda(M; \mathbf{x}; \mathbf{b})$ for the DOAS ratio using the simplified 1D RT in (24) as functions of (τ_{O_2}, M), that is, the two variables sampled respectively by OCI and MAS. From left to right: $\mathbf{x} = \{p_{top}, \Delta p\}$ in (27)–(28), and $\mathbf{b} = \{\tau_p\}$ in (29). We see from (27) that the iso-curves in the leftmost panel are exactly hyperbolic

- First, we consider a *hyper*-spectral sensor with 5-nm resolution that modulates the DOAS ratio, via *contiguous sampling of* λ in $\tau_{O_2}(\lambda)$, which appears three times in (24). At that resolution, it takes at most four spectral pixels to completely cover the A-band. OCI, the minimum payload for the PACE mission, is our inspiration.
- Second, we consider a *multi*-spectral (discrete sampling of λ) MAS that samples the A-band with *only two channels*, one in-band and one in the continuum. Such an instrument modulates the DOAS ratio via $|\mu|$, the (absolute) cosine of the VZA, in the airmass factor M. M also three times in (24), two of which are in the product $M(\mu_0, |\mu|)\tau_{O_2}(\lambda)$. The sensor on the proposed Multi-Angle Imager for Aerosols (MAIA) investigation (Diner et al. 2018) at NASA will have such non-polarimetric A-band channels, and it is indeed our inspiration. Notwithstanding, a growing number of space-based MASs will have polarimetric capability, some including A-band channels, others not.

Finally, we will consider the capability that results from a fused OCI-and-MAS sensor pair that may be brought together as part of a future mission.

The upper-left panel in Fig. 5 shows the same spectral information as in Fig. 1 but displayed as a DOAS ratio using only the first term in (24), i.e., assuming a reflective surface rather than a scattering aerosol layer. We therefore set $p_{\text{top}} = p_0$, $\Delta p = 0$, and take $M = 3$, as a typical value, e.g., 60° SZA and nadir viewing. Also in this panel, we have delineated above the top axis an example of where the boundaries between spectral channels could be in the OCI, with 5-nm resolution. Below the bottom axis is another indicator of the position and width for the "in-band" channel for MAIA.

Imagine now a sequence of three or four spectral pixels, each 5-nm wide, that can fall randomly on the essentially "line-by-line" signal in the upper-left panel and average it to that degraded resolution. One instance is indicated at the top of the upper-left panel. The outcomes are plotted in the upper-right panel of Fig. 5 as a function of the short-wavelength edge of the shortest wavelength pixel using, for simplicity, a square bandpass filter 5 nm wide. Anticipated DOAS ratio error bars of 2% are also plotted. For the purposes of passive aerosol profiling, there are good, bad and average scenarios color-coded respectively in green (one instance), red (two instances) and orange (one instance). In a "bad" situation, two of the three spectral pixels give the same ratio, to within the noise. A "good" situation is when there are three different ratios as broadly separated as possible, not counting a near-unity value. An "average" or "random" situation is anything else. In the remainder, we will assume the average case highlighted in this panel. Positions of those four spectral pixels are reproduced along the upper axis of the upper-left panel. There is hardly any absorption in the 4th one. In principle, it could be adopted as the continuum "reference" channel, and used to form the three other DOAS ratios. In practice, reference DOAS channels are located at a small but finite spectral distance from the absorption band.

As our sixth and final modeling simplification, we will use a single M-dependent "effective" value of τ_{O_2} for the spectral pixel,[5] defined as

---

[5]This is in lieu of spectrally averaging the numerator of the DOAS ratio in (24), as well as the Jacobians from (27)–(29), for every different choice of the aerosol and geometry parameters. This is
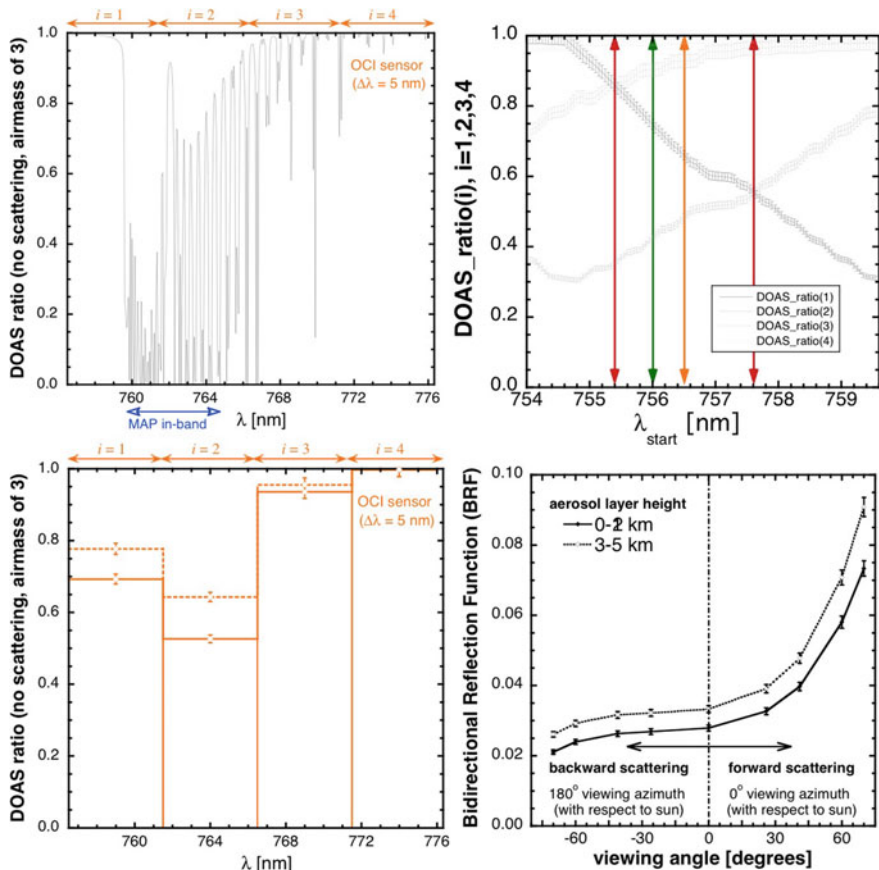
Fig. 5 In the left-side panels, we see O_2 A-band spectroscopy in DOAS ratio format at quasi-line-by-line (top) and OCI (bottom) resolutions, respectively; further descriptions in main text. The upper-right panel shows the diversity of coarse-resolution DOAS ratios that can be obtained depending on the spectral registration of the OCI instrument, i.e., exactly where the hyper-spectral channels (with 5-nm bandpass) are separated in the A-band region; colored vertical lines explained in main text. Lower-right panel shows synthetic BRFs for the single "in-band" spectral channel of a notional MAS with the same sampling of VZAs as MISR/Terra. In the two lower panels, hyper-spectral (left) and multi-angle (right) signatures are shown for two aerosol profiles: 0–2 km and 3–5 km, over a dark surface; otherwise identical AOT of 0.1 and smoke-like optical properties, as described in main text. Expected instrument error bars are plotted, 2% for DOAS ratios and 3% for BRFs, showing that the two profiles can be distinguished, especially in the multi-angle observations

$$\tau_{O_2}^{(i)}(M) = -\frac{1}{M} \ln \left(\frac{1}{\Delta \lambda_i} \int_{\Delta \lambda_i} e^{-M \tau_{O_2}(\lambda)} d\lambda \right), \tag{30}$$

---

tantamount to driving the monochromatic forward model in (24) with a single "effective" wavelength $\lambda_i(M)$ such that $\tau_{O_2}(\lambda_i(M)) = \tau_{O_2}^{(i)}(M)$ for $i = 1, 2, 3$.

where $M = 3$ (e.g., 60° SZA with nadir viewing) and $i = 1, 2, 3$. Thus, the dominant first term in (24) would be accurately estimated for thin ($\Delta p \ll p_0$) low-level ($p_{top} \approx p_0$) aerosols. Specifically, this leads to $\tau_{O_2}^{(1)}(3) = 0.142$, $\tau_{O_2}^{(2)}(3) = 0.243$, and $\tau_{O_2}^{(3)}(3) = 0.027$ for our random (i.e., neither "lucky" nor "unlucky") spectral registration scenario in Fig. 5, upper-right panel.

In the lower-left panel, we see the piecewise constant DOAS ratios for the four spectral observations. Here again, 2% error bars are plotted. Two aerosol profiles are used. In both cases, we take AOT $\tau_p = 0.1$. In one case (solid lines), the aerosol layer is between the surface and the top of the planetary boundary layer set at 2 km ($\{p_{top}, \Delta p\} = \{789,224\}$ hPa); this is a frequent occurrence. In the other scenario (dashed lines), the aerosol is assumed to be lofted by synoptic winds to 3–5 km altitude ($\{p_{top}, \Delta p\} = \{542,154\}$ hPa); this also happens frequently. In particular, it can be Saharan dust spreading clear across the Atlantic Ocean. At any rate, the two profiles can be distinguished (accounting for instrument noise levels), at least in the two shortest wavelength ranges.

Finally, the lower-right panel of Fig. 5 shows BRFs from a notional MAS sensor for the "in-band" spectral channel indicated near the lower axis of the upper-left panel, also treated with an effective $\tau_{O_2}^{MAS}(3) = 0.384$. In this case, 3% error bars are used, as is standard practice in absolute radiometry. The nine angles used here are chosen to be identical to those of the Multi-angle Imaging Spectro-Radiometer (MISR) (Diner et al. 1998): nadir, ±26.1°, ±45.6°, ±60.0°, and ±70.5°, with "+" meaning "fore" and "−" meaning "aft" looks that correspond respectively to forward and backward scattering in Terra's descending sun-synchronous orbit in the northern hemisphere, and conversely south of the equator. A smooth smoke-like phase function[6] was used, along with a relatively low single scattering albedo (SSA) of 0.88071, in (21)–(22). The same two aerosol profiles as in the lower-left panel were used, and we see that they are even more distinguishable in this synthetic MAS observation, even in view of the 3% radiometric noise level.

5 Information Content Analyses

5.1 IC Quantification Method

We will continue to denote our forward model for the DOAS signal in (24) as $\mathbf{F}(\mathbf{x}; \mathbf{b})$ where $\mathbf{x} = (p_{top}, \Delta p)^T$, superscript "T" means transpose, is our A-band remote sensing target. The number of unknowns in the atmospheric state space is therefore $m = 2$. The only quantity that is required is $\mathbf{b} = \tau_p$, the non-retrieved aerosol property that is estimated (with known uncertainty) using other spectroscopic and/or multi-angle polarimetric signals. Then $\mathbf{F}(\mathbf{x}, \mathbf{b}) = (F_1(\mathbf{x}, \mathbf{b}), \dots, F_n(\mathbf{x}, \mathbf{b}))^T$ is an n-dimensional

---

[6]Spherical (Mie) particles following a lognormal size distribution, with $r_g = 0.06$ μm geometric mean radius, $\ln(\sigma_g) = 0.6$, and complex refractive index $1.518 - 0.02368$ i at $\lambda = 446$ nm; although this is not the O_2 A-band wavelength, the resulting phase function is representative.

vector in measurement space containing all the observed DOAS ratios using either/or mono-angle spectroscopic and multi-angle bi-spectral sensing since both methods can modulate the key quantity in the DOAS ratio in (24), namely, $M(\mu_0, |\mu|)\tau_{O_2}(\lambda)$.

We will use concepts from Rodgers' OE theory (Rodgers 2000), which is summarized in Appendix "IC Analysis in Optimal Estimation Theory" . That framework is heavily dependent on the $n \times m$ Jacobian matrices $\mathbf{J} = \partial \mathbf{F}/\partial \mathbf{x}$ and $\mathbf{J_b} = \partial \mathbf{F}/\partial \mathbf{b}$, which are evaluated in closed form for our simplified 1D RT model in a previous Section.

To use the formalism of OE, we need to specify numerically these three variance/covariance matrices:

- $\mathbf{S_y}$, for observation error on DOAS ratios, is taken as diagonal (vanishing covariances), assuming a relative 1.5% error (one standard deviation or "StDev"), which is $\approx 1/2$ of the conventional assumption for absolute radiometric error of $\approx 3\%$, but is typical for ratios of radiances from the same sensor since different channels tend to drift in the similar ways; we therefore have $\sqrt{S_{yii}} = 0.015 \times F_i(\mathbf{x}; \mathbf{b})$, $i = 1, \ldots, n$.
- $\mathbf{S_b}$, for the uncertainty on $\mathbf{b} = \tau_p$, is taken as $\sqrt{S_{b11}} = \max\{0.02, 0.15\tau_p\}$, which is commensurate with quantities used in operational space-based aerosol remote sensing (Kahn et al. 2010).[7]
- $\mathbf{S_a}$, a priori uncertainty on retrieved properties $\mathbf{x} = (p_{top}, \Delta p)^T$, is also assumed diagonal, with $\sqrt{S_{a11}} = 250$ hPa and $\sqrt{S_{a22}} = 150$ hPa; in essence, this means that the top of the aerosol layer is log-normally distributed with a 68% chance of laying between ≈ 1 and ≈ 7 km, and that uncertainty on its thickness is $\approx 1200/p_{top}$ [km], with p_{top} expressed in hPa, e.g., approximately ± 1.3 km for a low (0–2 km) aerosol layer and ± 2 km for a lofted (3–5 km) one. These are quite loose constraints.

To be consistent with the single-scattering model, we set $\tau_p = 0.1$ at the A-band, which translates roughly to AOT of ≈ 0.15 at 550 nm, assuming an Angstrom exponent α of about unity, and ≈ 0.2 if α increases (smaller particles) to about 2. That choice leads to $\sqrt{S_{b11}} = 0.02$.

Our goal is to quantify the retrievability of the aerosol profile defined parametrically in Sect. 3. To this effect, we invoke the notion of partial Degree Of Freedom (DOF) A_{jj} for the jth retrieved retrieved property x_j that is defined in (48). As explained in Appendix "IC Analysis in Optimal Estimation Theory", it is closely related to "information content" per se in the formalism of OE. It is actually a measure of information *gain* going from the prior to posterior uncertainty on $\mathbf{x}$; see Appendix "IC Analysis in Optimal Estimation Theory". In the remainder, we assign $x_1 = p_{top}$ and $x_2 = \Delta p$. To estimate the anticipated retrieval error from A_{jj}, we use (49).

---

[7]In all cases discussed in this article, the eigenvalues of $\mathbf{J_b S_b J_b^T}$ (forward modeling error) are very small compared to the diagonal terms in $\mathbf{S_y}$ (measurement error).
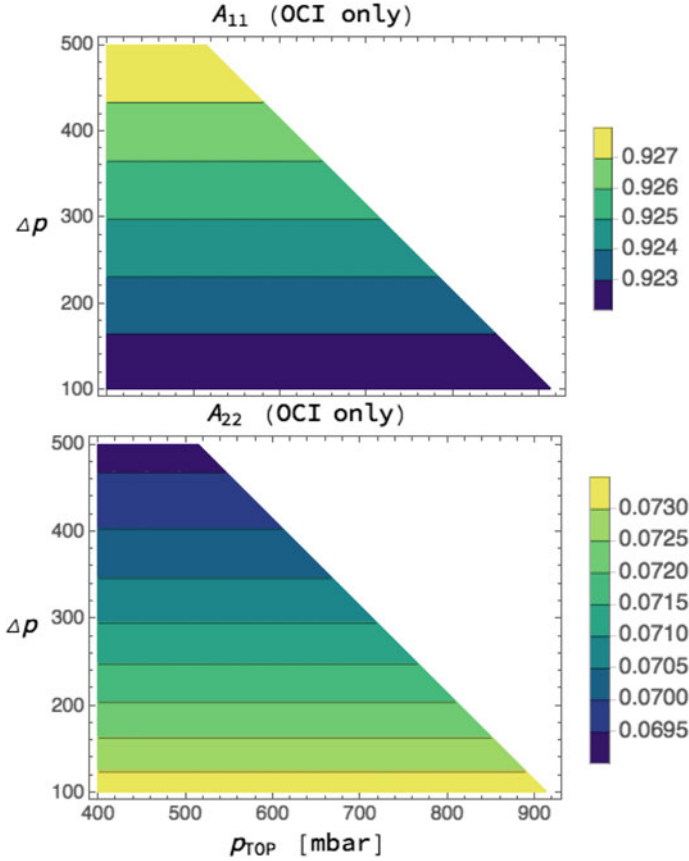
Fig. 6 Partial DOFs for p_{top} (left panel) and Δp (right panel) as functions of $(p_{\text{top}}, \Delta p)$, for the OCI instrument alone. Detailed discussion in main text

5.2 OCI

What happens if we only have the hyper-spectral OCI sensor registered spectrally in an average configuration with respect to the A-band, as defined in Sect. 4? In this case, we have $n = 3$, with one DOAS ratio for each of the three in-band spectral pixels for a single direction (assumed, for simplicity, to be nadir). Figure 6 shows A_{11} for p_{top} and A_{22} for Δp based on this minimal PACE instrumentation, as functions of $\mathbf{x}^{\text{T}} = (p_{\text{top}}, \Delta p)$. We see that the two extremes of DOF are occurring.

The partial DOF is indeed relatively high for p_{top}, hence excellent chances of retrievability with a typical error StDev[p_{top}] ≈ 68 hPa ($\approx 27\%$ of the 250 hPa prior), from (49) using a mid-range DOF. That translates approximately to 0.6 Km for aerosols in the boundary layer, and ≈ 0.9 km when lofted. At least for small AOT, A_{11} (hence StDev[p_{top}]) is insensitive to p_{top} itself, and there is a slight but systematic increase (decrease) with an increase in Δp.

In sharp contrast, we have very low partial DOF for Δp, hence no chance of retrievability, and we see from (49) that the uncertainty is still almost at the prior of ± 150 hPa, i.e., the observations did not help much. We can trace this to the fact that one cannot separate dependencies on p_{top} and Δp in (24) at 1st order in $M\tau_{O_2}$, cf. (26). The 2nd and higher-order terms are required, and they are much smaller in magnitude.

5.3 MAS

What happens if we can only use the MAS? We start with an "in-band" channel with an effective $\tau_{O_2}^{MAS} = 0.384$ (same center but slightly narrower than shown in Fig. 5). We will assume $n = 5$, which is the anticipated minimum number of directions to be collected by the proposed MAIA sensor for any given target, which will give us a lower bound on the A-band performance of that instrument. We thus obtain one DOAS ratio for each of the following VZAs: nadir, $\pm 30°$, and $\pm 60°$. That translates to $|\mu|_j = 1$, $\sqrt{3}/2$ (twice), and $1/2$ (twice), hence $M_j = 3$, 3.155 (twice), and 4 (twice) for a SZA of $60°$. Alternatively, we can again take $n = 3$, and use a weighting matrix $\mathbf{W} = \text{Diag}[1, 2, 2])$ for the sum-of-squared-residuals term, hence $\mathbf{W}^{-1} = \text{Diag}[1, 1/2, 1/2]$ that multiplies $\mathbf{S_y}$ in (34).

Computations of the partial DOFs for p_{top} and Δp as functions of the unknown aerosol profile parameters $(p_{top}, \Delta p)$ yield 2D plots (not shown) that look very similar to the left-hand panel in Fig. 6, even if it is about Δp—only the ranges of values in the legend change. We have listed the top-to-bottom DOF trends in the 2nd data column of Table 1, labeled "MAS #1," where we also display $\text{StDev}[p_{top}]$ and its counterpart for z_{top} with proper physical units. We see that, with respect to OCI, there is a modest

Table 1 DOFs for p_{top} and Δp, and StDev's for p_{top} and z_{top}, using either OCI, the notional MAS for PACE (#1), both, or the enhanced MAS planned for MAIA (#2). DOFs are described by the top ($\Delta p = 500$ hPa) to bottom ($\Delta p = 100$ hPa) trend in plots such as in Fig. 6. For Δp with OCI, we note (with *italics*) that the very slight trend is in the opposite direction. $\text{StDev}[p_{top}]$ is from (49), and $\text{StDev}[z_{top}] \approx H_{mol} \times \text{StDev}[p_{top}]/(p_{top} + \Delta p/2)$, with $H_{mol} = 8$ km for the mid-range DOF. Aerosols described as either "lofted layer" or "boundary layer" refer respectively to $(p_{top}, \Delta p) = (542,154)$ hPa, i.e., 3–5 km, and (789,224) hPa, i.e., 0–2 km

	Sensor:	OCI	MAS #1	both	MAS #2
DOF	Top/thicker	0.929	0.945	0.949	0.949
p_{top}	Bottom/thinner	0.922	0.931	0.931	0.931
StDev of	Top/thicker	67	59	56	56
p_{top} (hPa)	Bottom/thinner	70	66	66	66
StDev of	Lofted layer	0.9	0.8	0.8	0.8
z_{top} (km)	Boundary layer	0.6	0.5	0.5	0.5
DOF	Top/thicker	*0.0690*	0.135	0.190	0.220
Δp	bottom/thinner	*0.0735*	0.085	0.090	0.110

increase in partial DOF for p_{top}, and a significant one for Δp. However, the latter improvement in partial DOF (A_{22} going from ≈ 0.071 to the range 0.085–0.135) is not enough to consider $x_2 = \Delta p$ as a reasonable target for retrieval.

An alternative MAS configuration was also tested where the in-band channel is exactly as shown in Fig. 5). That filter choice (6 nm FWHM, up from 5 nm) leads to $\tau_{\text{O}_2}^{\text{MAS}} = 0.363$, somewhat lower than previously considered due to the admittance of wavelengths closer to the A-band's wings, but has an improved SNR. However, the number and range of VZAs is now set to the maxima considered for MAIA: nadir, $\pm 26°$, $\pm 43°$, $\pm 56°$, and $\pm 67°$. The outcome in DOFs and in retrieved parameter uncertainties is entered in the last column of Table 1, labeled "MAS #2." As expected, the improved VZA sampling more than compensates the decrease in effective $\tau_{\text{O}_2}^{\text{MAS}}$, as we see improved performance—still not enough however to go after Δp as a target.

5.4 OCI Plus MAS

What happens when we fuse the OCI- and MAS#1-based DOAS ratios into a single dual-modality observation? The outcome for DOFs and StDevs is displayed in the next-to-last column of Table 1, labeled "both." We see that, as expected, retrieval performance is improved over OCI or MAS#1 alone on both accounts, p_{top} and Δp. However, the enhancement is small for p_{top} as both sensors already do well. In short, OCI takes us most of the way for p_{top}, even with an "average" spectral sampling in the sense defined previously, in connection with Fig. 5. By coincidence, the p_{top} DOF range for OCI+MAS#1 matches (to 3 significant digits) that of the enhanced MAS#2 (last column of Table 1), identical retrieval uncertainties thus follow.

On the other hand, the increase in IC for Δp from adding the MAS#1 to OCI is dramatic (DOF can more than double). However, it does not reach the level achieved by MAS#2, and thus remains too small to justify keeping Δp as a remote sensing target. It should therefore be moved from $\mathbf{x}$ to $\mathbf{b}$ in the arguments of the forward model $\mathbf{F}(\mathbf{x}; \mathbf{b})$, with a prescribed value and an uncertainty either set to $\text{StDev}_a[\Delta p]$ for OCI, or obtained from Table 1 for MAS and when both sensors are used.

6 Summary & Outlook

A new challenge in passive atmospheric remote sensing from space is to determine not only the column-integrated aerosol load, namely, aerosol optical thickness (AOT), but also the height of the aerosol layer, especially if it is a strongly absorbing species such as dust, smoke or volcanic ash. These aerosol types can indeed reside virtually anywhere in the atmospheric column depending on injection height, synoptic winds and distance from sources.

Scientific motivation for this new thrust in aerosol profiling comes from air quality (where the interest is in particulates near the surface, as opposed to the column-integrated AOT), climate physics (e.g., aerosol impacts on clouds), atmospheric correction (in particular, for advanced ocean color studies), and other application areas. Technological motivation for not relying entirely on active lidar systems stems from: (*i*) their higher cost and complexity compared to passive systems; (*ii*) their limited spatial sampling in the horizontal compared to passive imaging sensors, even though their vertical resolution is far superior than any passive method can achieve; and (*iii*) their limited capability in terms of AOT for detection of aerosol layer base as well as top. It is widely acknowledged that, in contrast with lidar, passive O_2 A-band profiling techniques work best for aerosol scenes with AOT$\gtrsim$0.3, which will often occur near sources and rarely elsewhere. Notwithstanding, we showed here that A-band aerosol profiling can be extended to lower AOTs, albeit at the cost of loosing one piece of information, namely, aerosol layer thickness, at least for the PACE-inspired sensor configurations that we studied. In short, active lidar and passive A-band methods for aerosol profiling are truly complementary.

In the Introduction, we surveyed two alternative techniques for inferring aerosol layer height and, possibly, thickness as well. They characteristically exploit multi-angle *imaging*. The first is based on basic geometry, kinematics and computer vision: feature tracking and stereoscopy inspired by MISR's operational cloud products (top height and motion vector), but implemented in an interactive software package that focuses on the detection and characterization of various kinds of aerosol plumes near their sources (where the general wind direction is clear even in a single view). The other is the emerging capability of aerosol plume tomography that is grounded in physics-based image synthesis using 3D radiative transfer (RT) in the single-scattering limit (also used here), and on computational techniques for solving large inverse problems. We refer the reader back to Sect. 1 for detailed references.

We used the definitions of information content (IC) and partial degrees of freedom (DOFs) from Rodgers' (2000) statistical theory of optimal estimation (OE) to investigate the feasibility of passive aerosol profiling using the oxygen A-band absorption feature at 759–772 nm. Being in the solar spectrum, O_2 A-band signals result from atmospheric scattering and surface reflection. We neglected the latter in the simplified RT model adopted here, which is justified for the ocean surface outside of the glint angles and shallow waters, and in the absence of high turbidity. Furthermore, scattering is limited to a single event in the aerosol layer, which is therefore assumed to be optically thin at A-band wavelengths, say, AOT(A-band) $\lesssim$0.1, hence AOT(550 nm) $\lesssim$0.15 to 0.2 depending on particle size. We exploit differential optical absorption spectroscopy (DOAS) ratios of radiances sampled both in-band and in the continuum, which are largely immune to *absolute* radiometric error; residual measurement error is assumed to be $\approx$1.5%. The resulting DOAS ratio estimation is independent of all inherent aerosol optical properties past AOT; specifically, its single scattering albedo and phase function (i.e., composition and size/shape) are irrelevant. In reality, they are simply second-order (Hollstein and Fischer 2014) since they impact only radiance scattered twice or more.

The motivation for these RT simplifications is that we ultimately want a closed-form analytical expression for the observed DOAS ratios and, consequently, of its derivatives ("Jacobians") with respect to aerosol properties, whether or not they are targets for A-band retrieval. Specifically, we considered a parametric representation of the aerosol profile based on the assumption of a single aerosol layer located precisely between two pressure levels. These levels are the A-band remote sensing unknowns. AOT is assumed to be retrieved using other spectral channels, with a known uncertainty.

Having in hand the desired closed-form expressions for the DOAS observations, we computed analytically all the elements of the associated Jacobian matrices, which are used extensively in IC analysis in the framework of OE theory. However, we first used these rarely available analytical Jacobians to successfully predict the outcome of the statistical IC analysis of DOAS signals using purely physical arguments. That said, our forward model based on such simplified RT is not expected to perform well in an actual retrieval algorithm due to systematic biases. It is only tasked with giving us yes/no answers to questions of retrievability of profile parameters for optically thin aerosol plumes, given a few different instrumental configurations.

Two future NASA efforts, the PACE (Plankton, Aerosol, Cloud, and ocean Ecosystem) mission (NASA 2017) and the proposed MAIA (Multi-Angle Imager for Aerosols) investigation (Liu and Diner 2017; Diner et al. 2018), dictated our choices of A-band DOAS signal sampling. In these cases, sampling is either hyper-spectral at a single viewing angle (PACE) or else a simple bi-spectral sampling of the A-band absorption feature at multiple viewing angles (MAIA). In short, both OE and simplified RT theories tell us that PACE's Ocean Color Instrument (OCI) data can support the retrieval of aerosol layer's height, but not its thickness, at least for low AOTs. The same conclusion is drawn for a MAIA-like multi-angle/multi-spectral sensor that does *not* use polarization in either of its in-band and reference A-band channels. To first order, OCI has very little sensitivity to the aerosol layer's thickness. Its multi-angle/multi-spectral counterpart has some thickness sensitivity, just not enough to actually retrieve it with any confidence, at least for the low AOTs and notional instrument configurations examined herein.

That said, the combination of polarization (Ding et al. 2016) and multi-angle capabilities is very promising (Kokhanovsky et al. 2015). The same can be said about far higher AOTs along with a forward model that accounts for multiple scattering since joint height-and-thickness retrievals have been demonstrated for clouds using either spectroscopy (Rozanov and Kokhanovsky 2004; Kokhanovsky and Rozanov 2005; Schuessler et al. 2014) or multi-angle imaging (Ferlay et al. 2010; Desmons et al. 2013; Merlin et al. 2016).

Future missions may have both types of instrument considered here, so it is of interest to see if fused OCI-like and multi-angle/multi-spectral data can improve the IC, which it does although still not enough to retrieve layer thickness, at least for the PACE-inspired instrument configurations examined here.

In closing, we have stated upfront our goal of delivering a tutorial that covers the physical and mathematical technicalities of this focused remote sensing research project. More specifically, either in the main text or in the Appendix, we spell out

definitions and sketch derivations in sufficient detail for advanced undergraduate and graduate STEM (Science, Technology, Engineering & Mathematics) students to see a broad transect of their basic training in math brought to bear on a single so-called "paper-study" in remote sensing retrieval feasibility. Such studies are increasingly viewed as a requirement before serious consideration of remote sensing hardware development. A (probably non-exhaustive) list of mathematical concepts encountered is: integral calculus in one and more dimensions, ordinary differential equations (ODEs) and their solution by satisfying given boundary conditions, multi-dimensional integral equations and their numerical solution as a Neumann series, multivariate Taylor series expansions, approximation by series truncation, partial derivatives, multivariate probability distributions and statistics, matrix algebra, nonlinear function minimization by conjugate gradient descent, ill-posed matrix inversion, conditional probabilities, maximum likelihood, and Shannon entropy.

Appendix: IC Analysis in Optimal Estimation Theory

Least-Squares Fit of a Forward Model to Data

The standard approach (Bevington and Robinson 1992; Press et al. 1989) to fitting a generally nonlinear forward model $\mathbf{F}(\mathbf{x}; \mathbf{b})$ to data $\mathbf{y}$ is to equate them, as n-dimensional vectors in measurement space, with allowance for random error $\boldsymbol{\epsilon}$:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}; \mathbf{b}) + \boldsymbol{\epsilon}. \tag{31}$$

where $\mathbf{x}$ is the m-dimensional vector in a state space that contains all the parameters used to find the best fit to the data; $\mathbf{b}$ is another vector in the space of non- or otherwise-retrieved parameters that are imperfectly known to within a known uncertainty.

We will consider two sources of error in $\boldsymbol{\epsilon}$: (*i*) instrumental error that affects $\mathbf{y}$, and (*ii*) forward model error that affects $\mathbf{F}(\mathbf{x}; \mathbf{b})$. Our main concern in the latter case is uncertainty on $\mathbf{b}$, the parts of a larger state vector that have to be treated as given when retrieving $\mathbf{x}$. Instrumental error is characterized by its $n \times n$ variance/covariance matrix

$$\mathbf{S_y} = \mathrm{E}[(\mathbf{y} - \mathrm{E}[\mathbf{y}])(\mathbf{y} - \mathrm{E}[\mathbf{y}])^{\mathrm{T}}] \tag{32}$$

where E[·] denotes mathematical expectation. Uncertainty on the non/otherwise-retrieved parameters is defined by their variance/co-variance matrix $\mathbf{S_b}$, which can be converted into the equivalent $\mathbf{S_F}$ of the measurement error matrix in (32) by using the Jacobian matrix $\mathbf{J_b} = \partial\mathbf{F}/\partial\mathbf{b}$ and its transpose $\mathbf{J_b^T}$:

$$\mathbf{S_F} = \mathbf{J_b}\mathbf{S_b}\mathbf{J_b^T}. \tag{33}$$

Since we are confident that measurement error in $\mathbf{y}$ and modeling error in $\mathbf{F(x; b)}$ are independent random variables, their (co)variance matrices just add to form:

$$\mathbf{S_\epsilon} = \mathbf{S_y} + \mathbf{S_F}. \tag{34}$$

The classic least-squares minimization method determines the estimate

$$\mathbf{x_{bp}} = \underset{\mathbf{x}}{\operatorname{argmin}}\left[\Psi_\mathbf{y}(\mathbf{x})\right] \tag{35}$$

for the *best possible* fit of the data for model $\mathbf{F(x; b)}$ by finding the minimum of the cost function

$$\Psi_\mathbf{y}(\mathbf{x}) = \frac{1}{2}(\mathbf{y} - \mathbf{F(x; b)})^\mathrm{T}\mathbf{S_\epsilon^{-1}}(\mathbf{y} - \mathbf{F(x; b)}). \tag{36}$$

In the frequent situation where $\mathbf{S_\epsilon}$ is diagonal, $\Psi_\mathbf{y}(\mathbf{x})$ is the sum of squared "observed–predicted" residuals down-weighted by the uncertainty on each measurement. Assuming, for the moment, that $\Psi_\mathbf{y}(\mathbf{x})$ is convex, one can use the iterative Gauss-Newton algorithm:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{S_x}\mathbf{J}^\mathrm{T}\mathbf{S_\epsilon^{-1}}(\mathbf{y} - \mathbf{F}(\mathbf{x}_i; \mathbf{b})), \tag{37}$$

where $\mathbf{J} = \partial\mathbf{F}/\partial\mathbf{x}$ is the usual $n \times m$ Jacobian matrix, and

$$\mathbf{S_x} = \left(\mathbf{J}^\mathrm{T}\mathbf{S_\epsilon^{-1}}\mathbf{J}\right)^{-1}, \tag{38}$$

using an arbitrary starting position $\mathbf{x}_0$; this method converges in one step if $\mathbf{F(x; b)}$ is linear in $\mathbf{x}$. Iteration is stopped when the non-dimensional weighted sum of residuals in (36), often denoted as $\chi^2/2$, is $\approx n/2$.

Note that $\mathbf{x_{bp}}$ from (35) will not be identical to the *true* state vector in (31) for at least two basic reasons, starting with the least serious:

- First, the data $\mathbf{y}$ are collected with a specific realization of the random instrument noise in ϵ, i.e., the component quantified by $\mathbf{S_y}$ in (34). Averaging over different measurements of $\mathbf{y}$ will reduce the noise level (by the square-root of the number of independent measurements) but not eliminate it. However, the distance between the true and estimated state vectors is likely to be on the order of the square-root of the diagonal elements of $\mathbf{S_x}$ in (38), i.e., the predicted retrieval error for each element of $\mathbf{x}$:

$$\mathrm{StDev}[x_j] = \sqrt{S_{\mathbf{x}jj}}\,(j = 1, \ldots, m). \tag{39}$$

- Moreover, most interesting problems, in remote sensing in particular, result in non-convex cost functions in (36), simply because $\mathbf{F}(\mathbf{x}; \mathbf{b})$ is non-linear. Hence there is no guarantee that it has a single global minimum to find by iterative Gauss-Newton steps. So, even assuming that the true state is near the global minimum of $\Psi_{\mathbf{y}}(\mathbf{x})$, depending on the initial guess $\mathbf{x}_0$ at the state vector, the iterative search could end in a local minimum that is generally not within the distance in (39) of the true state. A well-known mitigation strategy for the non-convexity problem is to use the Levenberg-Marquardt minimization algorithm (Marquardt 1963; Bevington and Robinson 1992) where $\mathbf{S}_{\mathbf{x}}^{-1} = \mathbf{J}^{\mathrm{T}}\mathbf{S}_{\epsilon}^{-1}\mathbf{J}$ in (38) is replaced by a weighted sum of itself and its diagonal elements. The weight starts high on the diagonal matrix, which make the iteration behave like a straightforward gradient descent, which can work for non-convex function minimization (away from inflection points). As the weight shifts away from the diagonal, the method reverts to Gauss-Newton steps as the correct minimum is approached in a state-space region where $\Psi_{\mathbf{y}}(\mathbf{x})$ is locally convex.

Finally, the whole inverse problem at hand can be ill-posed, that is, $\mathbf{S}_{\mathbf{x}}^{-1}$ can be nearly singular, thus making its inversion in (38) highly unstable. The trajectory of iterative minimization in (37) will then be very sensitive to small perturbations of $\mathbf{y}$, e.g., instrumental noise. We now introduce a solution to this last issue since in remote sensing there is often much redundancy in the observations with respect to the unknowns in the retrieval.

Optimal Estimation

Rodgers' (2000) theory of optimal estimation (OE) revisits the above inverse problem of parameter estimation in the forward signal model from a Bayesian perspective, as a means for dealing with chronic ill-posedness by introducing regularization. Bayes theorem relates two conditional probabilities and two unconditional counterparts:

$$P(\mathbf{x}|\mathbf{y}) = P(\mathbf{y}|\mathbf{x})P(\mathbf{x})/P(\mathbf{y}), \tag{40}$$

where the last one is just a normalization factor of no particular interest here. $P(\mathbf{x})$ is the "prior" or "a priori" probability, while $P(\mathbf{x}|\mathbf{y})$ is the "posterior" probability.

Letting $|\cdot|$ denote the determinant of a square matrix, and assuming gaussian distributions, we have

$$\log P(\mathbf{x}|\mathbf{y}) = -\frac{1}{2}\left[(\mathbf{x} - \hat{\mathbf{x}})^{\mathrm{T}}\mathbf{S}_{\mathbf{x}}^{-1}(\mathbf{x} - \hat{\mathbf{x}}) + \log|\mathbf{S}_{\mathbf{x}}| + m\log(2\pi)\right] \tag{41}$$

for the posterior probability of atmospheric state properties $\mathbf{x}$, given data $\mathbf{y}$. Also, we note that $\hat{\mathbf{x}}$ is a *new* estimate of the prevailing state vector. Both $\hat{\mathbf{x}}$ and $\mathbf{S}_{\mathbf{x}}$ will depend on $\mathbf{y}$, $\mathbf{F}(\mathbf{x}; \mathbf{b})$, and related quantities, as shown below. In addition, we have

$$\log P(\mathbf{x}) = -\frac{1}{2} \left[(\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) + \log |\mathbf{S}_a| + m \log(2\pi) \right] \qquad (42)$$

for the a priori probability of atmospheric state, i.e., what we know about it before any observations are made. Finally, we have

$$\log P(\mathbf{y}|\mathbf{x}) = -\Psi_{\mathbf{y}}(\mathbf{x}) - \frac{1}{2} \log \left[(2\pi)^n |\mathbf{S}_\epsilon| \right] \qquad (43)$$

for the probability of obtaining the specific observations $\mathbf{y}$, given the atmospheric state $\mathbf{x}$, i.e., probability of seeing the random residuals used in (36) to estimate the cost function $\Psi_{\mathbf{y}}(\mathbf{x})$.

One then defines $\hat{\mathbf{x}}$ as the state with maximum likelihood, i.e., that maximizes $P(\mathbf{y}|\mathbf{x})P(\mathbf{x})$, a product of two gaussian PDFs. In other words,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left[\Psi_{\mathbf{y}}(\mathbf{x}) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) \right], \qquad (44)$$

instead of (35). The Gauss-Newton algorithm in (37) still applies but, rather than (38), we now have

$$\mathbf{S}_{\mathbf{x}} = \left(\mathbf{J}^T \mathbf{S}_\epsilon^{-1} \mathbf{J} + \mathbf{S}_a^{-1} \right)^{-1}. \qquad (45)$$

Although there may be better choices, one can always start the iterations with $\mathbf{x}_0 = \mathbf{x}_a$. As in (39), the diagonal elements $S_{\mathbf{x}jj}$ of $\mathbf{S}_{\mathbf{x}}$ are the *posterior* estimates of variance on the retrieved properties in $\hat{\mathbf{x}}$.

Degrees of Freedom

At any rate, and contrary to (38), the matrix inversion problem in (45) is by design well-posed, thanks to the presence of $\mathbf{S}_a^{-1}$. However, convergence goes to $\hat{\mathbf{x}}$ rather than $\mathbf{x}_{bp}$ in (35).

It is therefore of interest to evaluate the $m \times m$ matrix

$$\mathbf{A} = \left(\mathbf{J}^T \mathbf{S}_\epsilon^{-1} \mathbf{J} + \mathbf{S}_a^{-1} \right)^{-1} \mathbf{J}^T \mathbf{S}_\epsilon^{-1} \mathbf{J}, \qquad (46)$$

i.e., the left-hand product of (45) with the potentially ill-conditioned matrix to be inverted in (38). Recall that these expressions are the predicted uncertainties on the retrieved state properties with and without regularization, respectively; in the latter case, however, the potentially unstable matrix inversion is not performed. A little algebra leads to the simpler expression

$$\mathbf{A} = \mathbf{I}_m - \mathbf{S}_{\mathbf{x}} \mathbf{S}_a^{-1}, \qquad (47)$$

Fig. 7 Schematic representation of an OE for $m = 2$ atmospheric state parameters relevant to the investigation presented in the main body. Ellipses are iso-likelihood contours where we have $(\mathbf{x} - E[\mathbf{x}])^T \mathbf{S}^{-1}(\mathbf{x} - E[\mathbf{x}]) = 1$, and these curves are traced respectively for the prior (grey) and posterior (black)
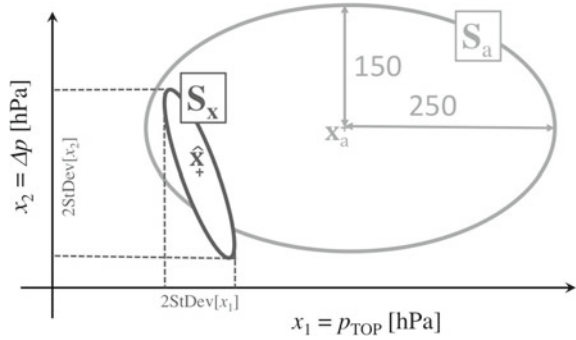
where $\mathbf{I}_m$ is the $m \times m$ identity matrix. Now, in OE theory, $\mathbf{A}$ is known as the "averaging kernel," and Rodgers (2000) shows that it can be defined simply as $\partial \hat{\mathbf{x}}/\partial \mathbf{x}$, where $\mathbf{x}$ is the true state parameter.

If $\mathbf{S}_a$ is diagonal (no prior covariances), then the diagonal terms of $\mathbf{A}$ are

$$A_{jj} = 1 - \frac{S_{\mathbf{x}jj}}{S_{ajj}} \tag{48}$$

for $j = 1, \ldots, m$. Since we can anticipate that $0 \leq S_{\mathbf{x}jj} \leq S_{ajj}$, we have $A_{jj} \in [0, 1]$. A_{jj} is known as the partial "degree of freedom" (DOF) for the retrieved property x_j. It is directly related to the predicted retrieval error

$$\text{StDev}[x_j] = \sqrt{S_{ajj} \times (1 - A_{jj})} = \text{StDev}_a[x_j]\sqrt{1 - A_{jj}}, \tag{49}$$

recalling that $\mathbf{S}_a$ is taken as diagonal.

Following the methodology of Merlin et al. (2016), we use A_{jj} and StDev$[x_j]$ from (48)–(49) extensively in the main body of this article. The former is an intuitive non-dimensional metric of the value added by the observations projected onto a specific state variable. If $S_{\mathbf{x}jj} \ll S_{ajj}$, then A_{jj} approaches unity, which means that the observations have vastly improved our knowledge of the state variable x_j. If, on the contrary, $S_{\mathbf{x}jj} \lesssim S_{ajj}$, then A_{jj} approaches zero, which means that the observations have not helped very much.

Figure 7 illustrates for $m = 2$ a typical concentration of probability measure (IC increase) in a familiar state space when going from the prior to posterior PDFs for $\mathbf{x}$. In both cases, we have traced, assuming gaussian PDFs, the lines of iso-probability value at $1/(2\pi)\sqrt{e|\mathbf{S}|}$, equivalently, where $(\mathbf{x} - E[\mathbf{x}])^T \mathbf{S}^{-1}(\mathbf{x} - E[\mathbf{x}]) = 1$ for easy visualization of the magnitudes of prior and posterior standard deviations. The area of each ellipsoid is $\propto |\mathbf{S}|$. More specifically, the regions inside the ellipses both account for 68% of the prior and posterior events. Note that, while $\mathbf{S}_a$ is diagonal (no covariance), $\mathbf{S}_\mathbf{x}$ is generally not.

A better-known quantification of overall retrieval performance in OE theory is

$$\text{Tr}[\mathbf{A}] = \sum_{j=1}^{m} A_{jj} \in [0, m], \tag{50}$$

which is the (implicitly, total) number of Degrees of Freedom. It approximates the maximum number of state properties (among m) that can be retrieved from the n observations in $\mathbf{y}$ in view of: (*i*) the instrumental noise ($\mathbf{S_y}$); (*ii*) forward model error ($\mathbf{S_F}$), in this case, from uncertainty in non-retrieved parameters ($\mathbf{S_b}$); (*iii*) sensitivities of the forward model to the state variables to be retrieved ($\mathbf{J}$), or not ($\mathbf{J_b}$); and (*iv*) prior information about the atmospheric state ($\mathbf{S_a}$).

The off-diagonal terms in $\mathbf{A}$ are non-dimensional quantities of the type $\text{E}\left[(x_1 - \hat{x}_1)(x_2 - \hat{x}_2)\right]$ divided by a diagonal element of $\mathbf{S_a}$, again assuming it is diagonal. These normalized cross-correlations describe how different state variables interfere (statistically) with one-another, two-by-two. Keeping these numbers small results in more regular (hence, more easily inverted) matrix $\mathbf{S_x^{-1}}$. That, in turn, will help the performance of the retrieval algorithm. In the design phase of the observation system, the forward model $\mathbf{F(x; b)}$ can thus be used to optimize the sampling of $\mathbf{y}$ in a way that keeps cross-variable interference as small as possible.

Entropy and Information Content

Lastly, we relate DOFs to information theoretical concepts, and thus justify our claim all along that we are quantifying Information Content per se. Following Shannon (1948), Rodgers (1998) defines the *increase* in information (equivalently, *decrease* in entropy) associated with the acquisition of observations $\mathbf{y}$ and their processing—by, e.g., OE methods—as

$$-\Delta H = -\frac{\log_2(|\mathbf{S_x}|) - \log_2(|\mathbf{S_a}|)}{2} = -\frac{1}{2}\log_2(|\mathbf{S_x S_a^{-1}}|) = -\frac{1}{2}\log_2(|\mathbf{I}_m - \mathbf{A}|), \tag{51}$$

when expressed in bits. Information gain $-\Delta H$ ranges from 0^+ ($|\mathbf{S_x}| \lesssim |\mathbf{S_a}|$) to ∞ ($|\mathbf{S_x}| \ll |\mathbf{S_a}|$). This follows directly from the expression for the entropy of a generic m-dimensional gaussian probability density function, $P_m(\text{E}[\mathbf{x}], \mathbf{S})$: from, e.g., (42), we have

$$H(m, |\mathbf{S}|) = -\text{E}[\log P_m(\text{E}[\mathbf{x}], \mathbf{S})] = \frac{1}{2}\log((\text{e}2\pi)^m |\mathbf{S}|), \tag{52}$$

which is naturally independent of the mean $\text{E}[\mathbf{x}]$. To visualize this IC increase (entropy decrease), $-\Delta H$ in (51) is $1/2$ the log of the ratio of the areas of the two ellipsoids in Fig. 7.

Abbreviations & Acronyms

1D	one-dimensional
AOT	aerosol optical thickness
BC	boundary condition
BRDF	Bi-directional Reflection Distribution Function
BRF	Bi-directional Reflection Factor
DOAS	Differential Optical Absorption Spectroscopy
DOF	Degree(s) Of Freedom
FWHM	Full-Width Half-Max
GSFC	Goddard Space Flight Center
IC	Information Content
IR	infra-red
JPL	Jet Propulsion Laboratory
MAIA	Multi-Angle Imager for Aerosols
MAS	Multi-Angle Sensor
NASA	National Aeronautics and Space Agency
OE	Optimal Estimation
OCI	Ocean Color Imager
PACE	Plankton, Aerosol, Cloud, and ocean Ecosystem
RT	radiative transfer
SNR	signal-to-noise ratio
SSA	single scattering albedo
StDev	Standard Deviation
SZA	solar zenith angle
TOA	Top-of-Atmosphere
UV	ultra-violet
VNIR	visible/near-infrared
VIS	visible
VZA	viewing zenith angle

References

Aides A, Schechner YY, Holodovsky V, Garay MJ, Davis AB (2013) Multi sky-view 3D aerosol distribution recovery. Opt Express 21(22):25820–25833

Bevington PR, Robinson DK (1992) Data reduction and error analysis for the physical sciences. WCB McGraw-Hill, New York (NY)

Boucher O, Randall D (2014) Chapter 7: clouds and aerosols. In: Cubasch U, Wuebbles D (eds) Climate change 2013: the physical science basis. Cambridge University Press, Cambridge (UK)

Chandrasekhar, S (1950) Radiative Transfer. Oxford University Press, Oxford (UK). [reprint by Dover Publications, New York (NY), 1960]

Corradini S, Cervino (2006)Aerosol extinction coefficient profile retrieval in the oxygen A-band considering multiple scattering atmosphere. Test case: SCIAMACHY nadir simulated measurements. J Quant Spectrosc Radiat Transfer 97(3):354–380

Davis AB, Kalashnikova OV, Diner DJ (2018) Aerosol layer height over water from the oxygen a-band: mono-angle spectroscopy and/or multi-angle radiometry. Remote Sens (in preparation)

Davis AB, Knyazikhin Y (2005) Chapter 3: A primer in 3D radiative transfer. In: Marshak A, Davis AB (eds) 3D radiative transfer in cloudy atmospheres. Springer, Heidelberg (Germany), pp 153–242

Davis AB, Polonsky IN, Marshak A (2009) Space-time Green functions for diffusive radiation transport, in application to active and passive cloud probing. In: Kokhanovsky AA (ed) Light scattering reviews, vol 4. Springer-Praxis, Heidelberg (Germany), pp 169–292

Desmons M, Ferlay N, Parol F, Mcharek L, Vanbauce C (2013) Improved information about the vertical location and extent of monolayer clouds from POLDER3 measurements in the oxygen A-band. Atmos Meas Tech 6(8):2221–2238

Diner DJ, Beckert JC, Reilly TH, Bruegge CJ, Conel JE, Kahn RA, Martonchik JV, Ackerman TP, Davies R, Gerstl SAW, Gordon HR, Muller J-P, Myneni RB, Sellers PJ, Pinty B, Verstraete MM (1998) Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. IEEE Trans Geosci Remote Sens 36(4):1072–1087

Diner DJ, Boland SW, Brauer M, Bruegge C, Burke KA, Chipman R, Di Girolamo L, Garay MJ, Hasheminassab S, Hyer E, Jerrett M, Jovanovic V, Kalashnikova OV, Liu Y, Lyapustin AI, Martin RV, Nastan A, Ostro BD, Ritz B, Schwartz J, Wang J, Xu F (2018) Advances in multiangle satellite remote sensing of speciated airborne particulate matter and association with adverse health effects: from MISR to MAIA. J Appl Remote Sens 12:042603

Ding S, Wang J, Xu X (2016) Polarimetric remote sensing in oxygen A and B bands: Sensitivity study and information content analysis for vertical profile of aerosols. Atmos Meas Tech 9(5):2077–2092

Dubuisson P, Frouin R, Dessailly D, Duforêt L, Léon J-F, Voss K, Antoine D (2009) Estimating the altitude of aerosol plumes over the ocean from reflectance ratio measurements in the O_2 A-band. Remote Sens Environ 113(9):1899–1911

Duforêt L, Frouin R, Dubuisson P (2007) Importance and estimation of aerosol vertical structure in satellite ocean-color remote sensing. Appl Opt 46(7):1107–1119

Evans KF, Marshak A (2005) Numerical methods. In: Marshak A, Davis A (eds) 3D radiative transfer in cloudy atmospheres, chapter 4. Springer, Heidelberg (Germany), pp 243–281

Ferlay N, Thieuleux F, Cornet C, Davis AB, Dubuisson P, Ducos F, Parol F, Riédi J, Vanbauce C (2010) Toward new inferences about cloud structures from multidirectional measurements in the oxygen A-band: middle-of-cloud pressure and cloud geometrical thickness from POLDER-3/PARASOL. J Appl Meteorol Climatol 49(12):2492–2507

Flower VJB, Kahn RA (2018) Karymsky volcano eruptive plume properties based on MISR multi-angle imagery and the volcanological implications. Atmos Chem Phys 18(6):3903–3918

Frankenberg C, Hasekamp O, O'Dell C, Sanghavi S, Butz A, Worden J (2012) Aerosol information content analysis of multi-angle high spectral resolution measurements and its benefit for high accuracy greenhouse gas retrievals. Atmos Meas Tech 5(7):1809–1821

Gabella M, Kisselev V, Perona G (1999) Retrieval of aerosol profile variations from reflected radiation in the oxygen absorption A band. Appl Opt 38(15):3190–3195

Garay MJ, Davis AB, Diner DJ (2016) Tomographic reconstruction of an aerosol plume using passive multiangle observations from the MISR satellite instrument. Geophys Res Lett 43(24):12590–12596

Hollstein A, Fischer J (2014) Retrieving aerosol height from the oxygen A band: a fast forward operator and sensitivity study concerning spectral resolution, instrumental noise, and surface inhomogeneity. Atmos Meas Techn 7(5):1429–1441

Kahn RA, Gaitley BJ, Garay MJ, Diner DJ, Eck TF, Smirnov A, Holben BN (2010) Multiangle imaging spectroradiometer global aerosol product assessment by comparison with the aerosol robotic network. J Geophys Res Atmos 115(D23):D23209

Kalashnikova OV, Garay MJ, Davis AB, Diner DJ, Martonchik JV (2011a) Sensitivity of multi-angle photo-polarimetry to vertical layering and mixing of absorbing aerosols: quantifying measurement uncertainties. J Quant Spectrosc Radiat Transfer 112(13):2149–2163

Kalashnikova OV, Garay MJ, Sokolik IN, Diner DJ, Kahn RA, Martonchik JV, Lee JN, Torres O, Yang W, Marshak A, Kassabian S, Chodas M (2011b) Capabilities and limitations of MISR aerosol products in dust-laden regions. Proc SPIE 8177:1–11

Kokhanovsky AA, Rozanov VV (2005) Cloud bottom altitude determination from a satellite. IEEE Geosci Remote Sens Lett 2(3):280–283

Kokhanovsky AA, Rozanov VV (2010) The determination of dust cloud altitudes from a satellite using hyperspectral measurements in the gaseous absorption band. Int J Remote Sens 31(10):2729–2744

Kokhanovsky AA, Davis AB, Cairns B, Dubovik O, Hasekamp OP, Sano I, Mukai S, Rozanov VV, Litvinov P, Lapyonok T, Kolomiets IS (2015) Space-based remote sensing of atmospheric aerosols: the multi-angle spectro-polarimetric frontier. Earth Sci Rev 145:85–116

Liu Y, Diner DJ (2017) Multi-angle imager for aerosols: a satellite investigation to benefit public health. Public Health Reports 132(1):14–17

Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 11:431–441

Merlin G, Riédi J, Labonnote LC, Cornet C, Davis AB, Dubuisson P, Desmons M, Ferlay N, Parol F (2016) Cloud information content analysis of multi-angular measurements in the oxygen A-band: application to 3MI and MSPI. Atmos Meas Tech 9(10):4977–4995

Min Q, Harrison LC (2004) Retrieval of atmospheric optical depth profiles from downward-looking high-resolution O_2 A-band measurements: optically thin conditions. J Atmos Sci 61(20):2469–2477

Mishchenko MI (2002) Vector radiative transfer equation for arbitrarily shaped and arbitrarily oriented particles: a microphysical derivation from statistical electromagnetics. Appl Opt 41:7114–7135

Moroney C, Davies R, Muller JP (2002) Operational retrieval of cloud-top heights using MISR data. IEEE Trans Geosci Remote Sens 40(7):1532–1540

NASA (2017) PACE: Plankton, Aerosol, Cloud, ocean Ecosystem. https://pace.gsfc.nasa.gov

National Academies of Sciences, Engineering, and Medicine (2018) Thriving on Our Changing Planet, A Decadal Strategy for Earth Observation from Space. The National Academies Press, Washington (DC). https://doi.org/10.17226/24938

Nelson DL, Chen Y, Kahn RA, Diner DJ, Mazzoni D (2008) Example applications of the MISR INteractive eXplorer (MINX) software tool to wildfire smoke plume analyses. Proc SPIE 7089:1–11

Nelson DL, Garay MJ, Kahn RA, Dunst BA (2013) Stereoscopic height and wind retrievals for aerosol plumes with the MISR INteractive eXplorer (MINX). Remote Sens 5(9):4593–4628

Pelletier B, Frouin R, Dubuisson P (2008) Retrieval of the aerosol vertical distribution from atmospheric radiance. Proc. SPIE 7150:7150R1–7150R9

Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1989) Numerical recipes. Cambridge University Press, Cambridge (UK)

Rodgers CD (1998) Information content and optimisation of high spectral resolution remote measurements. Adv Space Res 21(3):361–367

Rodgers CD (2000) Inverse methods for atmospheric sounding: theory and practice. World Scientific, Singapore

Rothman LS (2010) The evolution and impact of the HITRAN molecular spectroscopic database. J Quant Spectrosc Radiat Transfer 111:1565–1567

Rozanov VV, Kokhanovsky AA (2004) Semianalytical cloud retrieval algorithm as applied to the cloud top altitude and the cloud geometrical thickness determination from top-of-atmosphere reflectance measurements in the oxygen A band. J Geophys Res Atmos 109(D5):D05202

Sanghavi S, Martonchik JV, Landgraf J, Platt U (2012) Retrieval of the optical depth and vertical distribution of particulate scatterers in the atmosphere using O_2 A- and B-band SCIAMACHY observations over kanpur: a case study. Atmos Meas Tech 5(5):1099–1119

Schuessler O, Rodriguez D, Loyola G, Doicu A, Spurr R (2014) Information content in the oxygen A-band for the retrieval of macrophysical cloud parameters. IEEE Trans Geosci Remote Sens 52(6):3246–3255

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423

Val Martin M, Logan JA, Kahn RA, Leung FY, Nelson DL, Diner DJ (2010) Smoke injection heights from fires in North America: Analysis of 5 years of satellite observations. Atmos Chem Phys 10(4):1491–1510

Vaughan MA, Powell KA, Winker DM, Hostetler CA, Kuehn RE, Hunt WH, Getzewich BJ, Young SA, Liu Z, McGill MJ (2009) Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements. J Atmos Ocean Technol 26(10):2034–2050

Winker DM, Pelon J, Coakley JA Jr, Ackerman SA, Charlson RJ, Colarco PR, Flamant P, Fu Q, Hoff RM, Kittaka C, Kubar TL, Le Treut H, McCormick MP, Mégie G, Poole L, Powell K, Trepte C, Vaughan MA, Wielicki BA (2010) The CALIPSO mission: a global 3D view of aerosols and clouds. Bull Am Meteorol Soc 91(9):1211–1230

Wu L, Hasekamp O, van Diedenhoven B, Cairns B, Yorks JE, Chowdhary J (2016). Passive remote sensing of aerosol layer height using near-UV multiangle polarization measurements. Geophys Res Lett 43(16):8783–8790. 2016GL069848

Xu X, Wang J, Wang Y, Zeng J, Torres O, Yang Y, Marshak A, Reid J, Miller S (2017) Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: first results from EPIC/DSCOVR at Lagrange-1 point. Geophys Res Lett 44(14):7544–7554

Optical Properties of Black Carbon Aggregates

Chao Liu

1 Introduction

Black carbon (BC, also widely referred to as soot), a typical carbonaceous aerosol, is an important by-product of incomplete combustion of fossil fuel, biomass, biofuel, etc. (Sorensen 2001; Bond and Bergstrom 2006; Bond et al. 2013; Shrestha et al. 2010; Sharma et al. 2013). As the most absorbing aerosol of solar radiation, BC directly warms atmosphere and reduces radiation reaching the surface, and, thus, plays a critical role on global and regional weather and climate (Jacobson 2001; Menon et al. 2002; Bond and Sun 2005; Ramanathan and Carmichael 2008; Schwarz et al. 2008; Chakrabarty et al. 2009; Scarnato et al. 2013; Bond et al. 2013).

BC can heat the atmosphere through absorbing solar radiation, and leads to positive direct radiative forcing (DRF) at the top of the atmosphere and negative DRF at the surface (Bergstrom et al. 2003; Bibi et al. 2017). BC's mass may be relatively small in the atmosphere compared with other aerosols, whereas it is the second strongest anthropogenic contributor to current global warming, after carbon dioxide (Ramanathan and Carmichael 2008). The annual BC emission is a few Tg per year with significant uncertainties from different studies (Bond et al. 2007; Ramanathan and Carmichael 2008), and the global average BC optical depth at 550 nm is approximately 0.01 (Chung et al. 2012a). Modeled BC radiative forcing ranges widely from ~0.1 to 1.2 W/m^2 at the top of the atmosphere with differences of almost an order of magnitude (Jacobson 2001; Ramanathan and Carmichael 2008; Bond et al. 2011, 2013), and observation constrained studies estimate total (natural + anthropogenic) BC forcing about 0.9–1.0 W/m^2 (Sato et al. 2003; Ramanathan and Carmichael 2008). The IPCC 5th Assessment Report (AR5) suggests a total radiative forcing of

C. Liu (✉)

School of Atmospheric Physics, Nanjing University of Information Science & Technology, Nanjing 210044, China
e-mail: chao_liu@nuist.edu.cn

atmospheric BC to be 0.6 W/m^2, with 0.4 W/m^2 of this from fossil fuel and biofuel (IPCC 2013).

The BC absorption can also exert the so-called semi-direct effect by enhancing cloud evaporation (Koch and Del Genio 2010). Furthermore, BC is responsible for indirect radiative forcing by acting as cloud condensation nuclei or ice nuclei, which affect cloud microphysical processes and change cloud lifetimes (Lohmann and Feichter 2005). In the atmosphere, the average lifetime of BC is a few days due to both wet removal and dry deposition, and it is much shorter than that of long-lived greenhouse gases. If BC particles deposit in cloud particles or snow and ice surface, it can not only dramatically change the cloud or surface reflectance, but also contribute to melting, in particular for Arctic sea ice (Conant et al. 2002; Painter et al. 2012; He et al. 2016, 2018).

BC is still one of the least known factors for DRF estimations, and significant efforts should be devoted to improve our understanding on BC optical properties. Observations may provide the most "reliable/realistic" aerosol properties, whereas both in situ and laboratory observations have their limitations on obtaining a complete and detailed knowledge of BC optical properties (Schmid et al. 2006; Mishchenko et al. 2007; Subramanian et al. 2007; Moosmüller et al. 2009; Smith et al. 2015). First, aerosols always exist in the atmosphere as mixtures of various kinds, and, to measure pure BC, instruments and experiments can hardly be designed to completely avoid the influence of other aerosol materials. Secondly, BC optical properties may be significantly influenced by observational instruments or sampling. For example, filter-based instruments (e.g., particle soot absorption photometer) collect aerosol particles in special filters, and this may break the geometry or size distribution of natural particles (thus affect BC optical properties) (Lawless et al. 2004; Arnott et al. 2005; Cappa et al. 2008). Furthermore, observations can only be carried out at limited conditions or wavelengths, and can not provide optical properties over the entire interested spectrum. As a result, numerical modeling becomes not only a meaningful addition to improve our understanding on BC optical properties, but also an important method for its radiative forcing simulation and optical-based property retrievals.

The BC optical properties have been extensively studied with numerous numerical models, either exact or approximate, and the importance of the complex BC aggregation structures on BC optical properties is well shown and widely noticed (Kahnert 2003; Kolokolova et al. 2006; Liu and Mishchenko 2005, 2007; Radney et al. 2014). This chapter presents a systematic numerical investigation on the scattering and absorption properties of BC with aggregation structures. We emphasize not only on the importance of nonsphericity on BC optical properties, but also on the capabilities and limitations of popular numerical models for the simulations. The optical properties, including the extinction, scattering, and absorption cross sections (C_{ext}, C_{sca}, and C_{abs}) or efficiencies (Q_{ext}, Q_{sca}, and Q_{abs}), single-scattering albedo (SSA), asymmetry factor (g), and angular-dependent scattering matrix elements (P_{ij}), will be extensively discussed in this chapter, and we consider most of them at a single visible wavelength of 500 nm. All results discussed in this chapter are for those of

randomly oriented particles. More details on the definitions of those optical quantities can be found in (Liou 2002; Mishchenko et al. 2002).

The rest of this chapter is organized as follows. The properties of BC aggregates, i.e., key parameters influencing optical properties, are introduced in Sect. 2. The capabilities and limitations of numerical models to simulate the optical properties of BC in the form of aggregates are discussed and compared in Sect. 3. Sections 4 and 5 introduce the optical properties of bare and mixed BC, respectively, and the influences of geometric parameters will be revealed in details. An interesting model to produce BC/aerosol colors in the ambient atmosphere is applied to show whether BC is black in Sect. 6. Section 7 concludes this study.

2 Black Carbon Aggregates

Before discussing the optical properties, BC geometry, size, and refractive index are briefly introduced and reviewed in this section, because they are key variables to determine the optical properties of any particles. Only by considering reliable BC properties, numerical investigation can obtain meaningful conclusions to improve our understanding on optical properties. Note that the review is not used to strictly limit the numerical simulations of this chapter, and our investigations in Sects. 4 and 5 may use variables beyond the observed values discussed here to maximize the advantages of numerical study, and to obtain a more complete understanding on BC optical properties.

2.1 Geometry

BC particles normally exist in the form of aggregates with hundreds or even thousands of small spheres in the range of 20–60 nm in diameter, which are commonly called spherule or monomer (Dankers and Leipertz 2004; Chakrabarty et al. 2006; Dukhin et al. 2007). Such geometries have been extensively proved by microscopic images (Gwaze et al. 2006; Kamimoto et al. 2007; Alexander et al. 2008; Tumolva et al. 2010; Wang et al. 2017). Once emitted into the atmosphere, BC particles undergo aging process, which can occur over a few hours after its emission and last for days (Lee 1983; Moffet and Prather 2009; Riemer et al. 2010; Peng et al. 2016; Liu et al. 2013, 2017b). BC aggregates collapse from lacy chain-like clusters into more compact ones during a short period of time (Martins et al. 1998; Posfai et al. 2004; Scarnato et al. 2013; Dlugach and Mishchenko 2015). Meanwhile, those aggregates can be mixed with other aerosol components by absorption or condensation of gaseous species, coagulation with other aerosols, and oxidation, and become inhomogeneous with coatings of nitrate, sulfate, or other carbonaceous materials.

Figure 1 illustrates some microscopic images of BC particles, examples of in situ or laboratory observations from different combustion fuels and conditions (Bambha
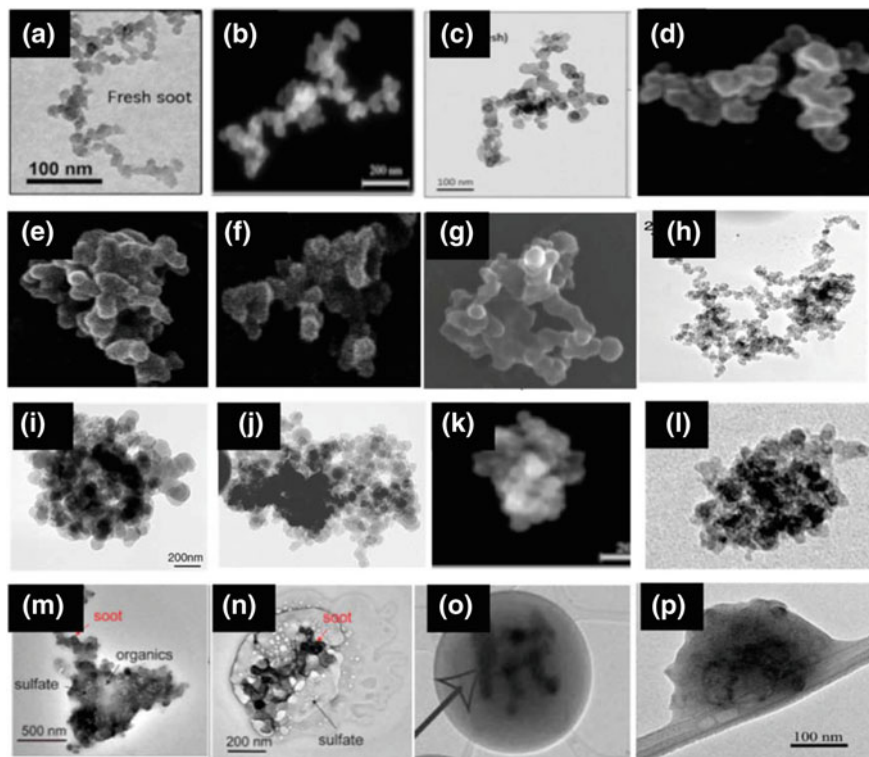
Fig. 1 Some microscopic images of black carbon particles from different in-situ or laboratory observations. **a** Ethylene-air flame from a laminar co-flow diffusion burner (Bambha and Michelsen 2015), **b** laboratory sample (Chen et al. 2016), **c** traffic emission (Pirjola et al. 2017), **d–f** non-premixed methane-air flame (Scarnato et al. 2013), **g** forest fire (field-emission from the Las Conchas fire) (China et al. 2013), **h–j** boundary layer sample at Sagres, Portugal (Li et al. 2003), **k** laboratory sample (Chen et al. 2016), **l** ethylene-air flame from a laminar co-flow diffusion burner (Bambha and Michelsen 2015), **m** Chinese urban (Wang et al. 2017), **n** mountaintop of Tai, China (Wang et al. 2017), **o** ambient sample from a coastal site (Freney et al. 2010), and **p** Mexico City (MC) plumes (Adachi et al. 2010)

and Michelsen 2015; Chen et al. 2016; Pirjola et al. 2017; Scarnato et al. 2013; China et al. 2013; Li et al. 2003; Wang et al. 2017; Freney et al. 2010; Adachi et al. 2010). Those particles show clear aggregation structures with small spherical monomers, and examples of both lacy and compacts particles are illustrated in the figure. By analyzing images like those in Fig. 1, the geometric parameters as well as the monomer properties can be retrieved. Figure 1 only shows the aggregate overall geometries, not the size, so the scale label is not included in all images.

To quantitatively describe the highly complex and randomly structured BC clusters, the concept of fractal aggregate (FA) shows great success and wide applications (Herd et al. 1992; Sorensen 2001). The FA assumes same-sized spheres to be point-

to-point attached with each other, and is mathematically described by the statistic scaling rule:

$$N = k_f \left(\frac{R_g}{a} \right)^{D_f}$$ (1)

where N is the number of monomers in an aggregate, and a is the monomer radius. k_f and D_f are the fractal prefactor and fractal dimension, respectively. R_g is the radius of gyration, given by:

$$R_g^2 = \frac{1}{N} \sum_{i=1}^{N} (\vec{r_i} - \vec{r_o})^2 + a^2$$ (2)

Here $\vec{r_i}$ is the position vector of the ith monomer, and $\vec{r_o}$ is that of the aggregate mass center. Note, different from the general expression, an additional a^2 term is included in Eq. (2), because the gyration radius is defined here as the mean square distance between mass center and all points on the spherule surfaces (Filippov et al. 2000). The definition is given because energy transfer or chemical reaction of aggregates with the carrier gas occurs on the surface. However, the definitions can only be significantly different for relatively small aggregates.

Following Eq. (1), as k_f or D_f increases, a relatively small R_g is required, which corresponds to a more compact particle. After originally emitted into the atmosphere, BC aggregates exhibit lacy structures with a small fractal dimension D_f, normally less than 2 (see Fig. 1a–d), whereas aged BC may be more compact (Fig. 1i–l, D_f maybe close to 3), and be mixed with other aerosols (Fig. 1m–p) (Sorenson 2001; Zhang et al. 2008; Chakrabarty et al. 2009; Wu et al. 2016).

In the framework of the FA, the detailed parameters of BC geometry have been extensively studied by analyzing the microscopic images, and various different fractal dimension and fractal prefactor are obtained. Meanwhile, the sizes of the monomers also gain a lot interests, and a slight size distribution on the monomer diameters is noticed (Sorensen et al. 1992; Sorensen and Roberts 1997; Sorensen 2001, 2011; Burr et al. 2012). Reported fractal dimension values range from 1.8 to 2.6 (Sorensen 2001; Kahnert and Devasthale 2011; Bambha and Michelsen 2015; Chen et al. 2016; Wang et al. 2017), and some BC images show aggregates close to solid spheres (Fig. 1l, o), the corresponding fractal dimension of which must be close to 3. Thus, we will consider fractal dimension values between 1.8 and 2.8 to account for variation on aggregate structure, and aggregates with D_f of 1.8 and 2.8 will be referred to as lacy and compact aggregates, respectively. The fractal prefactor is less discussed in the literatures due to its relatively less influence on aggregate optical properties, which will be noticed in Sect. 3, and a value of 1.2 estimated by Sorensen and Reborts (1997) will be used as the default value in this study.

To accurately calculate the optical properties of a given particle, its geometries should be rigorously presented. With the control of the statistic scaling rule, it is possible to generate those FAs numerically, and there are various different algorithms
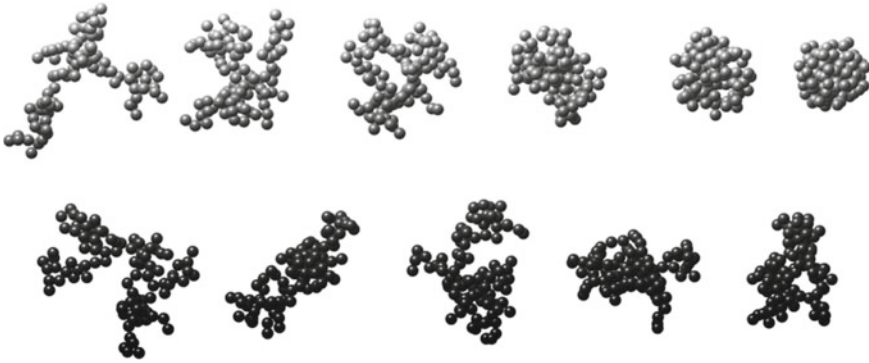
Fig. 2 Examples of numerically generated fractal aggregates with $N = 100$. The upper panels are aggregates with a fixed fractal prefactor of 1.2 and fractal dimensions from 1.8 to 2.8 in steps of 0.2, and the lower panels are those with a fixed fractal dimension of 1.8, and fractal prefactors from 1.0 to 1.8 in steps of 0.2

for this (Witten and Sander 1981; Jullien and Kolb 1984; Oh and Sorensen 1998; Filippove et al. 2000; Sorensen 2001). Diffusion-limited aggregation model, allowing monomers to aggregate in the space by random walking, is one of the first numerical models to produce aggregates with a complex fractal structure (Witten and Sander 1981; Oh and Sorensen 1998), and has been developed to various different versions. However, the fractal parameters cannot be prescribed in diffusion-limited models, and similar models include the chemical-limited algorithm (Jullien and Kolb 1984). The tunable aggregation model allows the fractal parameters to be arbitrarily defined, and is more flexible for aggregate generation (Filippove et al. 2000). Among them, the particle-cluster based models give particles by adding monomers one by one, and a particle grow larger with a quite clear "center". This limitation can be avoided by the cluster-cluster based aggregation algorithm. In the cluster-cluster algorithm, two original aggregates that follow the statistic scaling rule themselves are aggregated to form a larger one that also satisfy the rule by giving an appropriate relative distance. No matter how the particles are generated, all the resulting particles should follow the statistic scaling rule. This study used a tunable particle-cluster aggregation algorithm, and the details for the aggregate formation can be found in Filippove et al. (2000).

Figure 2 shows some examples of numerically generated BC aggregates with 100 monomers. The upper panels are aggregates with a fixed fractal prefactor of 1.2 and fractal dimensions from 1.8 to 2.8 in steps of 0.2, and the lower panels are those with a fixed fractal dimension of 1.8, and fractal prefactors from 1.0 to 1.8 in steps of 0.2. With the increase of D_f, aggregates change from lacy clusters to compact ones, and an aggregate with D_f of 2.8 is almost a sphere (similar to Fig. 1i–l). Comparing Figs. 1 and 2, it is safe to conclude that the FA is a reasonable model to represent realistic BC geometries, and, by changing the fractal parameters, it provides sufficient flexibility to define aggregates of different kinds.

2.2 Size

As one of the most widely observed quantities for aerosol studies, particle size distribution is a relatively well known BC microphysical property. Due to the irregular geometries and the different principles used for size measurements, various different definitions have been applied for BC size, e.g., mobility diameter, mass equivalent diameter, and size based on projected areas (Bond et al. 2002; Schnaiter et al. 2005; Chakrabarty et al. 2006; Kirchstetter and Novakov 2007; Reddington et al. 2013; Wang et al. 2015). For example, the Scanning Mobility Particle Sizer (SMPS) measures the mobility diameter, which is also sensitive to particle shape. The Single Particle Soot Photometer (SP2), a particle-by-particle based instrument that incandesces non-BC materials, gives the size of BC core (Gao et al. 2007). The projected areas derived from BC microscopic images are also an important variable to express BC sizes. The SP2 instruments give BC volume-equivalent diameters from a few tens nanometer to almost 1 μm (Reddington et al. 2013; Wang et al. 2015). Similar ranges of BC size are obtained from the equivalent projected area diameters (Chakrabarty et al. 2006, 2014) and the SMPS (Schnaiter et al. 2005; Kirchstetter and Novakov 2007), although the definitions on the sizes as well as the distributions are quite different. To simplify and unify the definition, this study uses two parameters interchangeably to represent BC size: number of monomers in an aggregate (N) and diameter (d_v) of equivalent volume sphere. For aggregates with same-sized monomers (radius of a), the two are related by: $d_v = 2a\sqrt[3]{N}$.

The lognormal size distributions are found to provide a reasonable representation on realistic BC aggregate and monomer size distributions, and are extensively used for numerical approximation of BC radiative properties and forcing (Bond et al. 2002; Chakrabarty et al. 2006; Moffet and Prather 2009; Chung et al. 2012a; Reddington et al. 2013; Wang et al. 2015). If the equivalent-volume diameter d_v is used to specify the lognormal size distribution, we have:

$$p(d_v) = \frac{1}{d_v\sqrt{2\pi}\ln\sigma}exp\left[-\left(\frac{\ln(d_v/d_{gmd})}{\sqrt{2}\ln\sigma}\right)^2\right] \tag{3}$$

The geometric mean diameter (GMD, i.e. d_{gmd}) in the lognormal size distribution indicates the 'averaged' particle size, and the geometric standard deviation (GSD, i.e. σ) is a shape parameter describing the spread of particle sizes. We use a GMD of 0.12 μm and a GSD of 1.5 as default BC size distribution in this study for bulk scattering properties (Alexander et al. 2008; Coz and Leck 2011; Reddington et al. 2013; Wang et al. 2015). As mentioned above, realistic BC monomers also show narrow size dispersion in the form of lognormal distribution (Köylü and Faeth 1992, 1994; Lehre et al. 2003; Dankers and Leipertz 2004; Chakrabarty et al. 2006, 2007; Liu et al. 2006; Bescond et al. 2014), and its effects on BC optical properties will be discussed in Sect. 4.2.
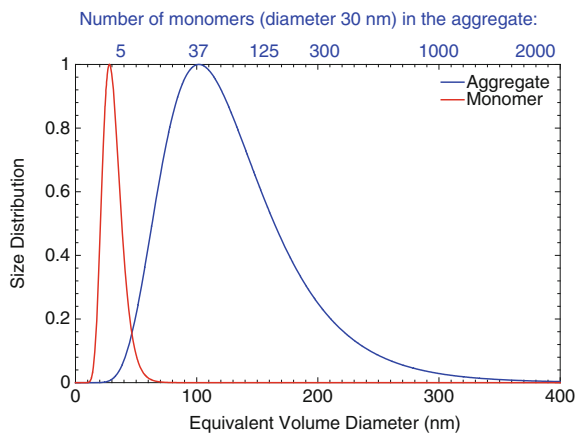
Fig. 3 Lognormal size distributions of black carbon aggregate and monomer. The sizes of non-spherical BC aggregates are defined by the diameters of corresponding volume equivalent sphere. The geometric mean diameters for aggregates and monomers are chosen to be 120 and 33 nm, and the geometric standard deviations are set to be 1.5 and 1.3, respectively

Figure 3 compares the lognormal size distributions for the BC aggregate and monomer, and diameters of the corresponding equivalent volume spheres are used to specify the aggregate size. The aggregate size distribution follows the default values mentioned above, i.e., $d_{gmd} = 0.12 \, \mu m$ and $\sigma = 1.5$, and the parameters are set to be 33 nm and 1.3 for BC monomers based on observations (Bond et al. 2002; Lehre et al. 2003; Bescond et al. 2014). We can see that the monomer diameters have a very narrow range of variation compared to aggregate volume-equivalent diameter distribution. The numbers of monomers in aggregate corresponding to the equivalent volume diameter are listed in the top for reference, and the number is calculated by assuming a monomer diameter of 30 nm, which is our default value. Besides Sect. 4.2 that discusses the influence of particle minor structures on BC optical properties, the rest of this chapter does not consider monomer size distributions, and studies only aggregates with same-sized monomers.

2.3 Refractive Index

The refractive index (RI), a wavelength dependent complex variable, is one of the most important parameters to determine particle optical properties. However, it is also one of the most uncertain properties of BC aerosols, because they cannot be directly measured. Current values of BC refractive indices are normally retrieved from observed absorption, scattering (or extinction) of suspended particles or from reflectance of compressed aerosol solids (Dalzell and Sarofim 1969; Chang and Charalampopoulos 1990; Stagg and Charalampopoulos 1993; Van-Hulle et al. 2002;

Moteki et al. 2010), whereas there are unneglectable uncertainties on both optical measurements and retrieval algorithms. Furthermore, BC materials from different combustions are possible to show different refractive indices, and this is verified by the previous studies (Hanel 1976; Sorensen 2001; Bergstrom et al. 2002; Bond and Bergstrom 2006; Lu et al. 2015). As a result, numerous datasets of BC refractive index have been developed over different spectrums to obtain its optical properties for radiative applications (D'Almeida et al. 1991; Krekov 1993; Hess et al. 1998). Some of those datasets extend refractive indices at particular wavelengths into those over continuum spectrum using the dispersion equations or Kramers-Kronig analysis (Dalzell and Sarofim 1969; Querry 1987; Chang and Charalampopoulos 1990; D'Almeida et al. 1991; Krekov 1993), so they can be applied for practical radiative applications conveniently. Those BC refractive indices are carefully reviewed and summarized in a couple of review researches (Sorensen 2001; Bond and Bergstrom 2006)

Figure 4 compares BC refractive indices from previous studies. Both real and imaginary parts of the BC particles show quite wide ranges of variations, and we eliminated results with real part larger than 2 and imaginary part much larger than 1.1. The real part varies between 1.5 and 2.0 with most values lying between 1.7 and 1.9. The imaginary part shows similar range of variation, and values from 0.4 up to 1.1 have been retrieved or used. Furthermore, quite different spectral variations over the wavelength are shown for both the real and imaginary parts. The figure clearly shows the uncertainty and large variation on BC refractive index. The refractive index of $1.8 + 0.6i$ at a wavelength of 500 nm will be used as our primary parameters for simulations, and exceptions to the aforementioned parameters will be specifically described in the discussion.
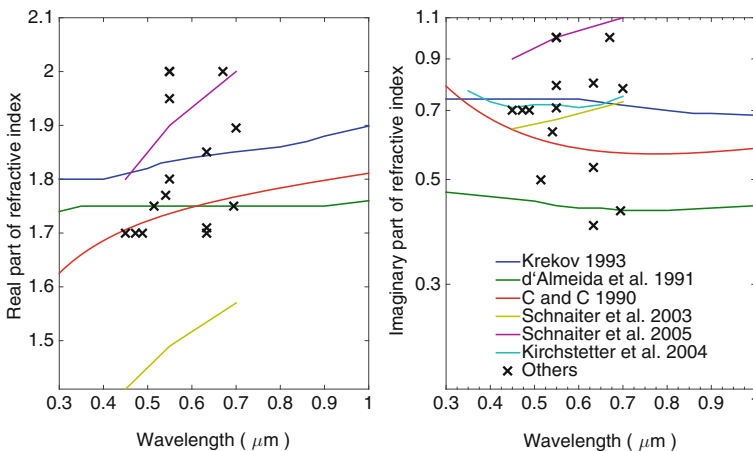
Fig. 4 The real and imaginary parts of black carbon refractive indices from various observational-based retrievals

Table 1 Parameters used to define BC aggregates for numerical simulations of light scattering properties

Property	Parameter	Default value	Range considered
Geometry	D_f	1.8	1.8–2.8
	k_f	1.2	1.0–1.8
	N	100 or 200	1–3000
	a	15 nm	5–50 nm
Size distribution	Aggregate	GMD: 120 nm; GSD: 1.5	50–200 nm
	Monomer	Same-sized: diameter of 30 nm Different-sized: GMD of 33 nm	5–50 nm
Refractive index	Real part	1.8	1.2–2.0
	Imaginary part	0.6	0.2–1.0

To summarize, this section provides a brief review on BC properties related to numerical simulations for its optical properties, and summarizes their ranges of variations for the optical property simulations. The default values as well as the ranges used for sensitive studies are listed in Table 1.

3 Numerical Methods for BC Optical Properties

3.1 Methods

Great efforts have been devoted to investigate the optical properties of atmospheric particles from numerical aspect. The scattering properties of particles much smaller than the incident wavelengths are normally approximated by the Rayleigh theory (Bohren and Huffman 2008; Moosmüller and Arnott 2009), and those of homogeneous or core-shell spheres are efficiently and accurately obtained by the Lorenz-Mie theory (Mie 1908; Bohren and Huffman 2008). However, atmospheric aerosols, such as BC particles studied in this chapter, are neither spherical nor much smaller than the incident wavelength (e.g., visible light), and optical properties given by corresponding equivalent spheres can hardly represent those of particles with aggregate structures (Li et al. 2010, 2016). The typical BC particles have equivalent-volume diameters in the same order of visible wavelength, e.g., a size parameter of 0.75 for a particle with a diameter of 0.12 μm at the wavelength of 0.5 μm, which makes the Rayleigh approximation challenged. However, BC monomers are normally in the Rayleigh regime, and, as one of the simplest and the most popular approximations, the Rayleigh-Debye-Gans approximation (RDG) is the very first description for the scattering and absorption of aggregates of those monomers (Sorensen et al. 1992; Farias et al. 1996b). The RDG represents only a first-order approximation, and the approximation is adopted by neglecting multiple scattering among the pri-

mary monomers that obey Rayleigh scattering themselves. It provides fairly good description of fractal soot scattering and absorption, whereas sometimes yields relative errors much larger than 15% (Sorensen et al. 1992; Cai et al. 1993; Farias et al. 1996b; Sorensen 2001; Van-Hulle et al. 2002; Sorensen et al. 2018). More details on the RDG model can be found in the classic review given by Sorensen (2001).

With the development of numerical methods for light scattering by non-spherical particles, more and more accurate methods are applied to account for the optical properties of BC aggregates. The Multiple Sphere T-Matrix method (MSTM) (Mackowski 1994, 2006, 2014; Mackowski and Mishchenko 1996, 2011; Liu et al. 2008) and the Generalized Multi-particle Mie (GMM) method (Xu 1995; Xu and Gustafson 2001; Xu and Khlebtsov 2003) are the most popular ones based on similar physical principles, and the two methods are specially designed to handle aggregates of spheres. Different from the RDG, the MSTM and GMM require the exact configurations of BC aggregates, i.e., the position, size, and component of each monomer. General methods for arbitrarily shaped particles, such as the discrete dipole approximation (DDA) (Draine and Flatau 1994; Yurkin and Hoekstra 2007, 2011), can also be applied to calculate aggregate optical properties, and they can be applied to further improve the complexity of BC geometries (Yon et al. 2015; Liu et al. 2016b). Besides, Liou et al. (2011) develop a geometric-optics approach coupled with surface-wave contributions for light scattering by aggregates, and the model shows great performance on approximating the optical properties of BC aggregates and snow grains.

Table 2 summarizes the popular numerical models discussed above to study BC optical properties, and all of them will be discussed in this study. We first demonstrate the significant errors that may be introduced by applying an approximate method (RDG or equivalent-sphere based approximations), and, then, compare the computational accuracy and efficiency of the three accurate models, namely the MSTM, GMM, and DDA. After this section, we use only those numerically accurate models to investigate BC optical properties. The MSTM is used as a benchmark model for idealized aggregates with perfect spherical monomers and no overlapping, and the DDA is used in Sect. 4.2 for irregular ones with minor imperfect structures.

3.2 Approximate Methods

Approximate methods such as the RDG do not require specified particle geometry for the simulation, whereas the MSTM calculates the optical properties of a given particle with precisely defined geometry. For a given set of geometric parameters, the fractal aggregates are randomly generated, and only the relative distances between the aggregates are fixed for the tunable aggregation algorithm. Each aggregate geometry can be referred to as an aggregate 'realization'. Before evaluation on the approximate method, we first verify how many aggregate realizations are needed for the accurate methods to give the optical properties of aggregates with certain structures.

Table 3 list ratios of the standard deviation to the mean value of optical properties of 20 aggregates with 10, 100, and 1000 monomers, and they indicate the relative

Table 2 Comparison of some popular numerical models used to calculate BC optical properties

Method	Description of the model	Source
MSTM	Accurate method for aggregates of spheres based on the superposition T-matrix method	Public code by Mackowski and Mishchenko (1996, 2011)
GMM	Accurate method for aggregates of spheres based on the interaction among spherical particles	Public code by Xu (1995)
DDA	Accurate model for particles with arbitrary shapes by solving Maxwell' equations in frequency domain	Public code by Yurkin and Hoestra (2011)
RDG	First-order approximation by neglecting multiple scattering among monomers	Self-developed following Sorensen (2001)
Equivalent Sphere	Use the optical properties of equivalent sphere to represent those of non-spherical particles	Public code by Bohren and Huffman (2008)

Table 3 Ratios of standard deviations to mean values of aggregate optical properties from different particle realizations. For the same aggregate parameters and size, optical properties of 20 aggregates with different realizations are calculated and averaged

D_f	N	C_{ext} (%)	C_{abs} (%)	C_{sca} (%)	g (%)
1.8	10	0.39	0.38	0.84	0.14
	100	0.43	0.30	3.5	1.7
	1000	0.21	0.13	1.8	0.15
2.8	10	0.37	0.28	0.37	0.0020
	100	0.18	0.19	0.18	0.0049
	1000	0.035	0.048	0.057	0.00046

variance of the optical properties due to aggregate realizations. The default variables are used for the simulations: a refractive index of $1.8+0.6i$ at a wavelength of $0.5\,\mu m$, and aggregates with a fractal prefactor of 1.2 and fractal dimensions of 1.8 and 2.8. Most ratios are under 0.5%, and only scattering cross sections for lacy aggregates have a ratio up to 3.5%. The ratios generally decrease with the increase of aggregate size. The optical property variations also become less significant for compact aggregates due to relatively spherical and compact structure. To summarize, a few aggregation realizations (less than 10) are enough to give the averaged optical properties of BC aggregates (Liu and Smallwood 2010; Liu et al. 2012b). However, for the computational efficiency, this study mostly uses results of a single aggregate realization to represent those of an ensemble with the same parameters.

Both spherical approximations based on the Lorenz-Mie method and those in the framework of the RDG introduce significant errors on calculating aggregate optical properties (Sorensen et al. 1992; Cai et al. 1993; Farias et al. 1996b; Li et al. 2010, 2016). Figure 5 compares the optical properties of BC particles
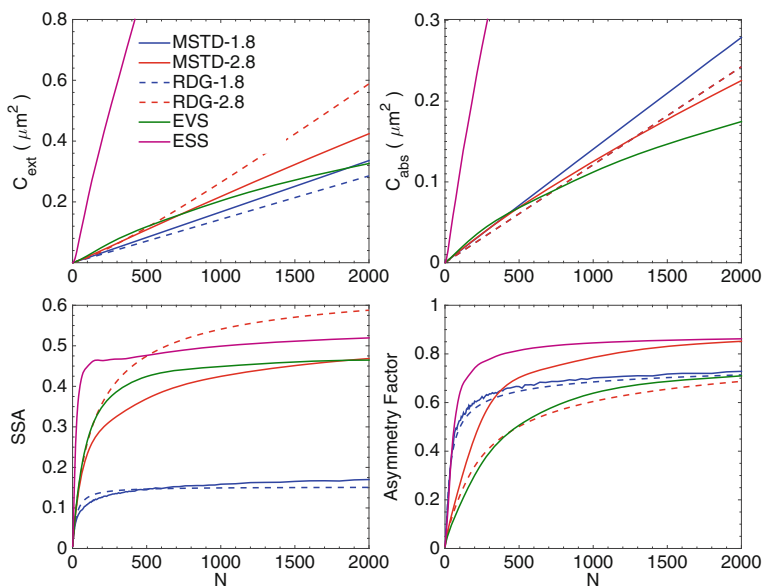
Fig. 5 Comparison of the optical properties (i.e., extinction (C_{ext}) and absorption (C_{abs}) cross sections, single-scattering albedo (SSA), and asymmetry factor) of black carbon aggregates given by different numerical models as functions of monomer number

calculated by numerical exact method (i.e. MSTM), the RDG approximation, and two equivalent spherical approximations. The extinction and absorption cross sections, single-scattering albedo, and the asymmetry factor as functions of aggregate size (i.e. monomer number in an aggregate) are illustrated in the figure. The corresponding results for equivalent volume and surface area spheres are referred to as the EVS and ESS, respectively. For each BC size (N), the MSTM results in Fig. 5 are averaged over those of 20 different aggregate realizations. We can see that the results given by the RDG and the equivalent spheres are significantly different from those given by the MSTM. For aggregates with $D_f = 1.8$, the RDG underestimates cross sections by approximately 13% to 25%, whereas shows asymmetry factors close to those given by the MSTM. As the aggregates become compact, the interaction among monomers becomes stronger, and, besides the absorption, the RDG results become worse. The relative errors of the extinction cross sections reach to almost 30%, and those for the SSA are even larger. However, the RDG approximates the absorption of compact aggregates relatively accurately with relative errors less than 10%. Note that the absorption given by the RDG is the sum of absorption of individual monomers given by the Rayleigh approximation, and is independent of aggregate structure. Because the RDG approximation gives the optical properties of given aggregates with really simple equations based on the Rayleigh approximation, its computational burden can be ignored. The EVS approximation can neither account for the effects of aggregation structures, and its performance varies at different conditions

Table 4 Bulk scattering properties averaged over an ensemble of black carbon particles with a lognormal size distribution

	C_{ext} (μm^2)	C_{abs} (μm^2)	C_{sca} (μm^2)	g
MSTM $D_f = 1.8$	0.021	0.018	0.0027	0.61
MSTM $D_f = 2.8$	0.026	0.018	0.0083	0.58
RDG $D_f = 1.8$	0.018	0.015	0.0024	0.58
RDG $D_f = 2.8$	0.026	0.015	0.010	0.46
EVS	0.029	0.019	0.010	0.41

for different quantities. The cross sections are relatively reasonable for aggregates with less than ~600 monomers, whereas the absorption is significantly underestimated for large aggregates. The ESS approximation completely overestimates the cross sections, because the sphere with the same surface as those of the aggregate has much larger particle volume.

Figure 5 indicates that the accuracy of the approximations for different quantities varies at different aggregate sizes, and Table 4 lists the bulk properties averaged over an ensemble of different-sized aggregates. The default BC size distribution with the GMD of 0.12 μm and the GSD of 1.5 is considered. With this size distribution, most aggregates have less than 300 monomers (see Fig. 3). All results listed are averaged values per single particle. Again, the RDG underestimates the extinction of lacy aggregates ($D_f = 1.8$) by 15%, mainly due to the underestimation on the absorption. Coincidentally, the RDG gives extinctions of compact aggregates close to the MSTM results, whereas underestimates and overestimates the absorption and scattering, respectively. The EVS approximation overestimates the cross sections by 5% (absorption) to over 200% (scattering of lacy aggregate). The ESS results deviate most significantly, and are not discussed.

The non-zero scattering matrix elements of aggregates with 400 monomers given by the numerically approximate and accurate methods are compared in Fig. 6. The RDG and MSTM show similar scattering phase functions for aggregates with lacy structure, whereas quite different results are noticed for the compact aggregates. The EVS phase function differs from both lacy and compact MSTM results. The ratios of other elements to the phase function for lacy aggregates features of the Rayleigh scattering due to relatively weak interactions among monomers, and only the MSTM results for compact aggregates and the EVS solutions show slight differences from those of the Rayleigh scattering.

Figure 7 is the same as Fig. 6 but for the scattering matrix elements of an ensemble of BC aggregates with the default size distribution. Although the aggregates with $D_f = 2.8$ are almost spherical and highly compact, their phase function still exhibits relative differences over 40% from those of the EVS at some scattering angles. The RDG results agree relatively closely to the MSTM ones for the lacy aggregates. Both the RDG and EVS approximations underestimate the forward scattering and overestimate the backward scattering. Again, the differences on other scattering matrix elements occurs only on compact BC given by the MSTM and the EVS.
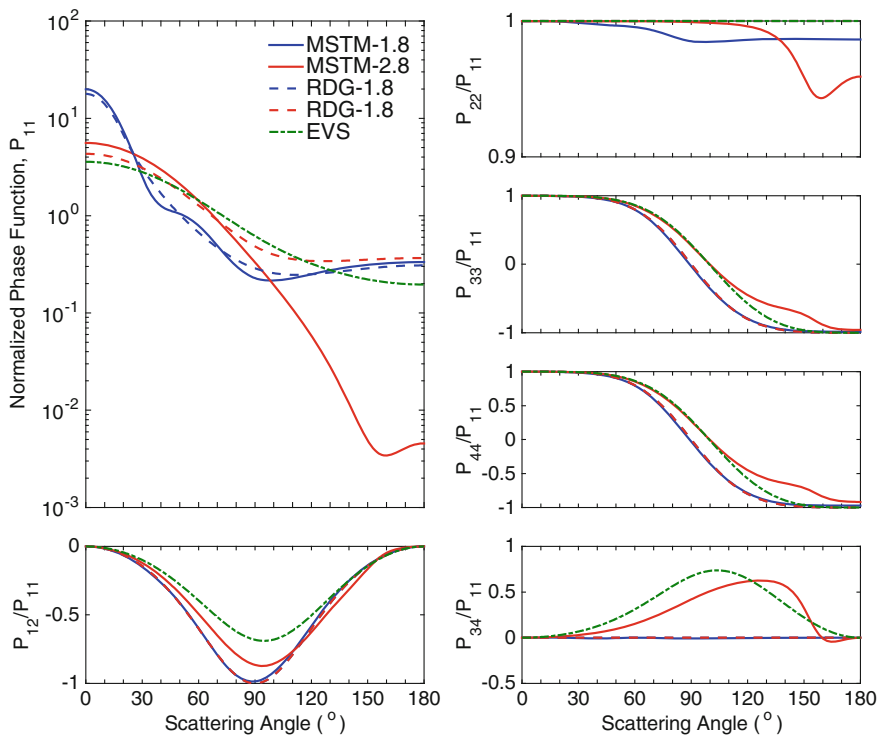
Fig. 6 Comparison of the non-zero scattering matrix elements of black carbon aggregates given by different numerical methods. The BC aggregate has 400 monomers, and the diameter of the equivalent volume sphere is 0.22 μm

The equivalent sphere approximation uses optical properties of spheres to represent those of non-spherical properties, and the "equivalence" can be defined for different quantities. Besides volume and surface area, radius of gyration is also a potential equivalent variable, and it is possible to have two equivalent parameters at the same time by using multiple spheres, i.e., both volume and ratio of volume to projected area. However, none of those spherical equivalent approximations gives reasonable solutions on the optical properties of BC aggregates (Li et al. 2010). Overall, we conclude that both the RDG and spherical approximation lead to numerical errors over 10% and up to over 100% for the optical properties of aggregates, although some model may show reasonable agreement on a particular quantity for a particular aggregate geometry. As a result, all results from here on will be calculated by the methods that can accurately account for the non-spherical geometries of BC aggregates.

To summarize, due to the different errors introduced by those approximate models, accurate methods and aggregation structure have to be applied to account for the optical properties of BC particles, although such computations are much more
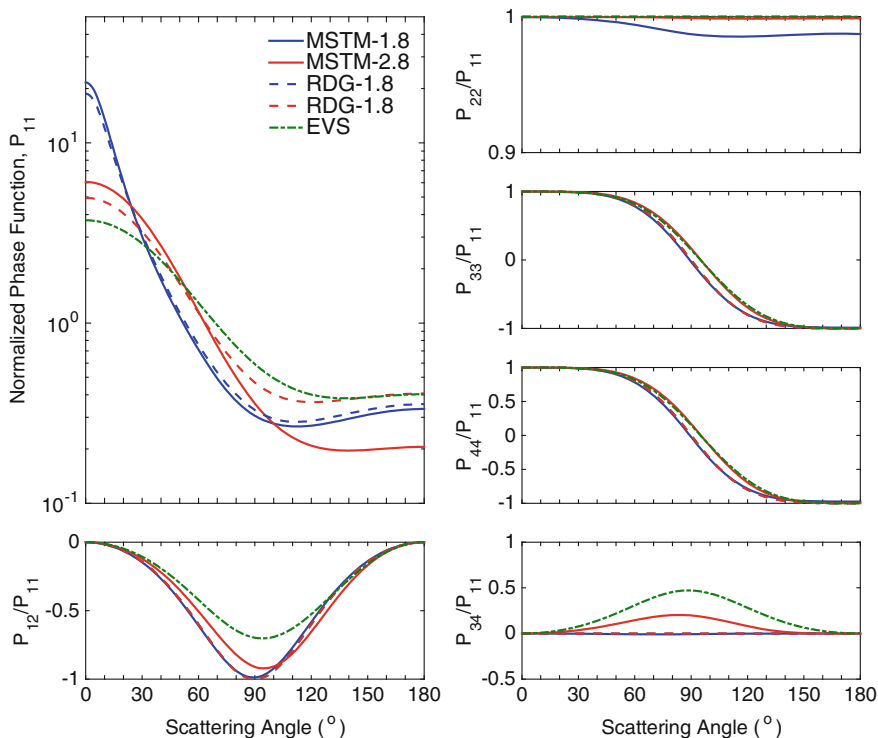
Fig. 7 Same as Fig. 6 but for those averaged over the given particle size distribution

complex and time consuming. Actually, even accurate methods show accuracy to different degrees, and will be investigated in the following subsection.

3.3 Accurate Methods

Among the various numerically accurate solutions, we focus on three, namely the MSTM, GMM, and DDA, because they are the most popular and appropriate ones widely used for BC applications. The MSTM and GMM are chosen because of their particular applicability and clear advantages on the accuracy and efficiency. The DDA is a general method for particles with arbitrary geometries, and has also been one of the most popular models for BC optical simulations (Draine and Flatau 1994; Kahnert et al. 2012; Scarnato et al. 2013; Moteki 2016), especially those with irregular aggregate structures. The time domain method, such as the finite-difference time domain (FDTD) method (Yang and Liou 1996) and the pseudo-spectral time domain (PSTD) method (Liu 1997), will not be considered in this study, because, for lacy particles such as aggregates, their computational efficiency is incomparable to

that of the DDA due to much larger computational domain. The DDA considers only dipoles inside the particle, whereas the entire domain including both the particle and the three-dimensional free space hosting the particle has to be computed in the time domain methods. Meanwhile, the monomer fine structure is not appropriate for the PSTD, which performs better with coarse spatial resolutions (Liu 1997; Liu et al. 2012a).

Both the MSTM and GMM consider a BC aggregate rigorously as a cluster of spheres defined by their positions, sizes, and refractive indices, whereas the DDA discretizes the spatial domain using small 'dipoles' to represent particles with arbitrary shapes. Thus, the discretization of DDA simulation becomes an additional parameter that should be considered, and it affects both the efficiency and accuracy of the simulation. Here, the size of dipole is defined by the number of dipoles in a wavelength (lambda), i.e., dpl. With the increase of dpl (i.e., smaller dipole size), the DDA represents the particle geometries better and results in more accurate optical properties, whereas more computational resources are required. Here, default convergence-controlling parameters in the DDA, MSTM, and GMM models are used for all simulations.

Figure 8 shows the accuracy and computational efficiency of the DDA with the increase of dpl, and the 'exact' solutions are given by the MSTM. We consider an aggregate with 200 monomers. With the increase of the dpl, the DDA results converge to those of the MSTM, whereas there are still relative errors of ~3% for the cross sections with dpl reaching up to 600. The DDA simulations normally suggest a dpl in the order of a few tens for particles much larger than the incident wavelength, and, here, a dpl value up to 600 seems still insufficient. Both the MSTM and DDA implementations used by this study are parallelized based on the MPI, and are carried out on the same single node with 24 64-bit 2.5 GHz processors. The bottom right panel shows the computational times used by the DDA and MSTM, and the MSTM becomes almost three orders of magnitude faster than the DDA simulations with a dpl of 600. For practical applications, the differences between the DDA and MSTM should be understood more carefully. The FA is an idealized geometry model with perfect spheres point-to-point attached, but realistic monomers do exhibit some degrees of imperfection, such as surface roughness and overlapping between monomers. The MSTM treats a particle as perfect spheres exactly the same as idealized FA, whereas the DDA can not represent the monomers perfectly as spheres in the discretized spatial domain. Thus, a portion of the DDA numerical error is introduced by such inaccurate shape representation, especially for simulations with a small dpl value. As a result, such 'error' can be understood as a correction for a more realistic representation of realistic particle geometries, although it lacks sufficient control for immediate application. Thus, additional caution should be taken on the choice of dpl when applying the DDA for BC optical property simulations, and normally a very large dpl, i.e., a few hundred, should be used to account for the small monomer structure. More details on the performance of the DDA for the particular application of BC aggregates can be found in Liu et al. (2018).

Figure 9 compares the integral scattering properties of aggregates with 200 monomers and different fractal dimensions given by the three numerically accu-
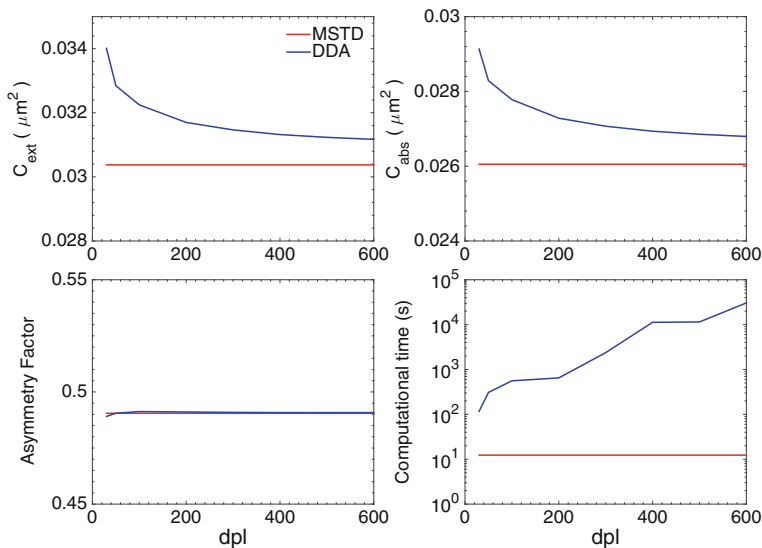
Fig. 8 Comparison of the optical properties (extinction and absorption cross sections, and asymmetry factor) and computational time given by the MSTM and DDA with different spatial resolutions for aggregates with 200 monomers. The spatial resolution of the DDA simulation is presented by the number of dipoles in a wavelength/lambda (dpl)

rate methods. The extinction and absorption cross sections, and asymmetry factors are compared, and the computational times used by the three methods are also given in the right bottom panel. The GMM code is serial, and its computational time is divided by the number of processes used by the other models for fair comparison. The GMM and MSTM give almost the same results with relative differences less than 1% for the cross sections, whereas those between the DDA and MSTM are around 4% (a dpl of 200 for the DDA simulations). The differences of the three methods on the asymmetry factors are relatively small, i.e., under 0.5% for most cases. For the computational efficiency, the MSTM is the most efficient one, which is an order of magnitude faster than the GMM. The efficiency of both the MSTM and GMM is less sensitive to aggregate geometry, whereas the DDA simulations become almost two orders of magnitude faster as the aggregates become compact.

The relative differences of the GMM and DDA results compared with the counterpart MSTM ones are listed in Table 5, and most of those results have been discussed in Fig. 9. Again, the GMM results are close to those of the MSTM, and relative differences of the DDA on the cross sections are still quite obvious at a dpl of 200. Meanwhile, both the GMM and DDA are less efficient than the MSTM.

Figure 10 illustrates the scattering matrix elements of aggregates with 200 monomers given by the MSTM, GMM, and DDA, and lacy ($D_f = 1.8$) and compact ($D_f = 2.8$) aggregates are considered. Although the three models give almost the same asymmetry factors (an integral of the scattering phase function), slight
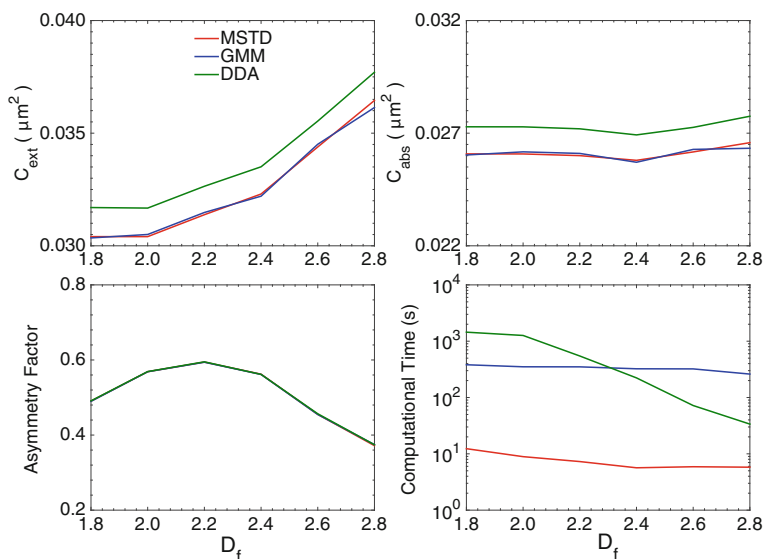
Fig. 9 Comparison of the optical properties of BC aggregates with different fractal dimensions given by the three methods. DDA simulations use a dpl of 200

Table 5 Relative differences of the optical properties of BC aggregates given by the GMM and DDA compared to the corresponding MSTM results, and their computational efficiency (ratios of computational time used by the GMM/DDA to that by the MSTM)

D_f	C_{ext} (%)		C_{abs} (%)		g (%)		Efficiency	
	GMM	DDA	GMM	DDA	GMM	DDA	GMM	DDA
1.8	−0.07	4.35	−0.07	4.72	−0.15	0.12	3.1×10^1	1.2×10^2
2.0	0.42	4.30	0.46	4.75	−0.02	0.05	3.9×10^1	1.4×10^2
2.2	0.42	4.12	0.48	4.68	−0.07	0.10	4.8×10^1	7.5×10^1
2.4	−0.19	3.88	−0.23	4.52	−0.14	−0.10	5.7×10^1	3.9×10^1
2.6	0.46	3.56	0.54	4.39	−0.07	0.36	5.5×10^1	1.2×10^1
2.8	−0.76	3.57	−0.83	4.54	0.72	0.67	4.5×10^1	5.8×10^0

differences are noticed on the phase functions of the lacy aggregates, especially in the backward direction. The agreement on the compact aggregate is much closer. Besides the P_{34} elements, almost the same results are obtained for other scattering matrix elements.

The DDA has its advantages for particles with irregular geometries, such as aggregates with non-spherical or overlapped monomers, so we would also present a comparison between the DDA and MSTM for an inhomogeneous case, i.e., aggregates with non-absorbing coating spheres, which can be handled by both methods. Table 6 compares the optical properties and computational times of the MSTM and DDA
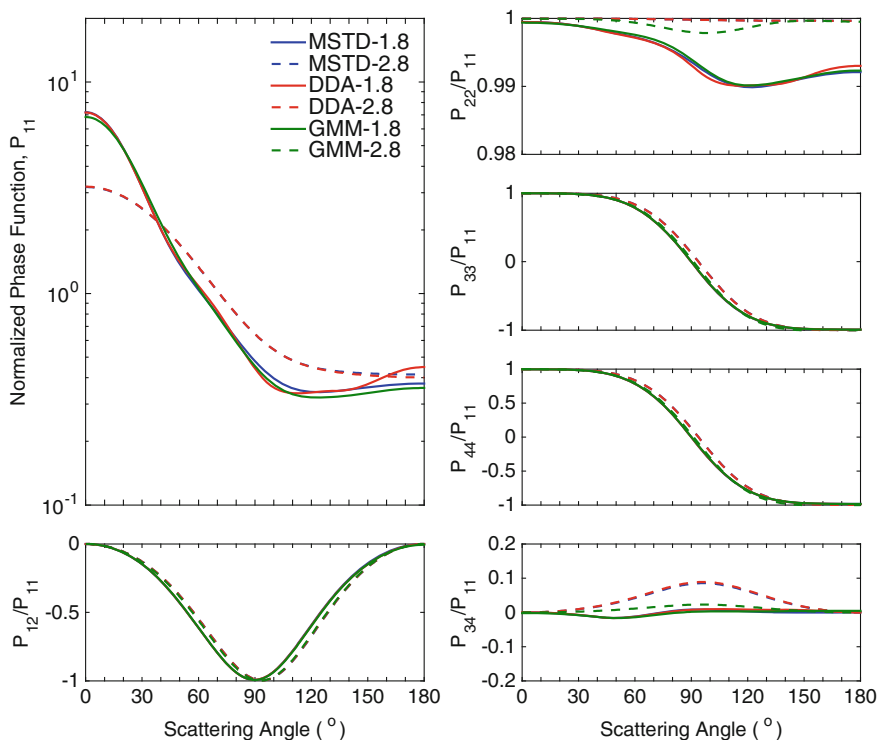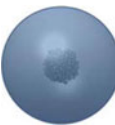
Fig. 10 Comparison of the scattering matrix elements of a black carbon aggregate with 200 monomers given by the MSTM, GMM and DDA

for bare and coated aggregates, and the geometries considered are illustrated in the table. Aggregates with 100 monomers are considered, and the blue sphere indicate the coating. The coating volume fraction $F_{coating}$ is set to be 0.99 to completely wrap the BC aggregates, and the coating refractive index is assumed to be 1.45. For the DDA simulations, the particles are discretized by grids with a dpl of 100 and 200. The accuracy and efficiency of the DDA for bare BC aggregates are similar to the results discussed above, whereas results for coated particles are quite different. First, with a large coating sphere, the accuracy of the DDA is substantially improved, and its results are almost the same as those of the MSTM at a dpl of either 100 or 200. Secondly, the MSTM becomes even more efficient for coated aggregates, whereas the DDA takes much more computational time because of larger particle overall size. Here, we just briefly compare the DDA and MSTM for coated aggregates from a methodological aspect, and more discussion on the optical properties of coated/mixed BC will be presented in Sect. 5.

With a better understanding on the numerical models and their performances, we can go forward to discuss the optical properties of BC aggregates, and the following two sections will present the optical of bare and coated BC aggregates respectively.

Table 6 Computational time of the MSTM and DDA methods for BC aggregates with a non-absorbing spherical coating

Particle	Method	Q_{ext}	Q_{abs}	SSA	g	Time
	MSTM	1.08	0.951	0.116	0.476	3.31
	DDA(100)	1.14(6%)	1.01(6%)	0.113(−3%)	0.477(0.2%)	69.9
	DDA(200)	1.12(4%)	0.995(5%)	0.114(−2%)	0.477(0.2%)	557
	MSTM	1.28	0.980	0.234	0.264	1.63
	DDA(100)	1.35(6%)	1.04(6%)	0.227(−3%)	0.264(0%)	8.45
	DDA(200)	1.33(4%)	1.02(4%)	0.228(−3%)	0.264(0%)	29.5
	MSTM	3.97	0.0776	0.980	0.756	1.82
	DDA(100)	3.97(0%)	0.0762(−2%)	0.981(0.1%)	0.756(0%)	475
	DDA(200)	3.97(0%)	0.0767(−1%)	0.981(0.1%)	0.756(0%)	4030
	MSTM	3.98	0.0694	0.983	0.753	1.20
	DDA(100)	3.98(0%)	0.0686(−1%)	0.983(0%)	0.754(0.1%)	407
	DDA(200)	3.98(0%)	0.0690(−1%)	0.983(0%)	0.753(0%)	3450

4 Optical Properties of Bare BC

The FA is the most popular and successful model to represent the geometries of natural BC particles, and we refer it as an overall, or first-order, description of BC particles. Meanwhile, realistic particles also show some detailed features, e.g., non-spherical or polydisperse monomers, and imperfect connection among monomers, which are not included in the framework of original FA, and these factors also influence the optical properties of bare BC (Dobbins and Megaridis 1991; Eggersdorfer and Pratsinis 2012; Skorupski and Mroczka 2014; Yon et al. 2015; Liu et al. 2015b). Those factors doesn't change BC overall geometry and influence the aggregates less significantly, and we refer them as minor, or imperfect, geometric features. This section discusses the influence of aggregate properties and minor imperfect structures on the optical properties of BC aggregates, and an optical property database based on the FA and accurate MSTM simulation is developed for practical BC related radiative transfer and retrieval applications.
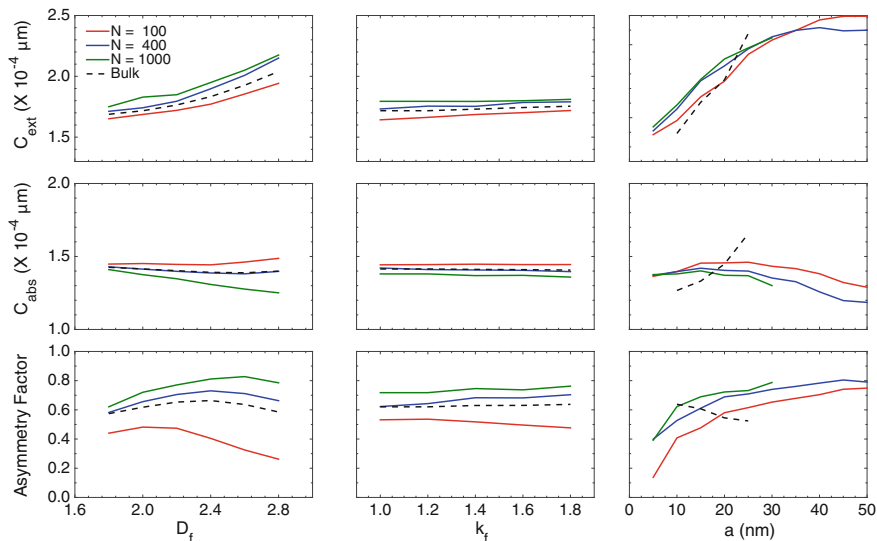
Fig. 11 Effects of fractal aggregate parameters, i.e., D_f (left), k_f (middle), and a (right), on the optical properties of black carbon aggregates. Both single-(colorful solid) and bulk-(black dashed) scattering properties are illustrated in the figure

4.1 Influence of Aggregate Properties

Figure 11 compares the optical properties of both single-sized ($N = 100$, 400, and 1000) and ensemble-averaged (dashed black curves) aggregates, and the x-axis variables from left to right panels are the fractal dimension, fractal prefactor, and monomer radius, respectively. To illustrate the properties in the same coordinate and to better understand the effects due to geometry, all cross sections are plotted as values for those per single monomer with a diameter of $a_o = 15$ nm. This is done by:

$$C_m = \frac{C_o}{N} \left(\frac{a_0}{a} \right)^3 \qquad (4)$$

Here, C_o is the original cross section of the given aggregate, and C_m is the normalized values per monomer illustrated in the figure. For example, the cross sections of the aggregate with 100 monomers plotted in the figure are the values divided by 100, and, for the right panels with different monomer radii, the differences on monomer volume are also normalized by that of the monomer with $a_o = 15$ nm by the $(a_o/a)^3$ term. For aggregates with different fractal dimensions, the extinction of highly compact aggregates is 20–30% larger than the lacy counterparts, mainly due to the significant increase in the scattering cross sections. The influence of fractal prefactor on the extinction is less significant with relative differences less than 5% as k_f varies between 1.0 and 1.8. Monomer size has the most significant influence on the extinction even

after the values normalized to those with the same volume. This is because that the scattering of a single monomer at this size range grow rapidly (e.g., proportional to a^6 based on the Rayleigh approximation). The effects of aggregate parameters on the absorption is less significant, and the absorption cross section per monomer is approximately 1.4×10^{-4} µm^2 (at refractive index of 1.8 + 0.6i). The variations of the asymmetry factors on the D_f, k_f, and a are not monotonic, and are quite different for aggregates with different sizes. The bulk properties averaged over the default size distribution are also given in the figure by the dashed black curves, and, for D_f and k_f in the left and middle panels, the bulk properties follow the same trends as those of single-sized aggregates. The bulk scattering properties of aggregates with different monomer diameters show different variations from those of single-sized aggregates. Because the size distribution is defined for equivalent volume diameter, and, as the monomer size increases, aggregates with smaller N are accounted to following the overall size distribution.

The effects of the fractal dimension on the scattering matrix elements of an ensemble of aggregates are shown in Fig. 12. With the increase of the fractal dimension, the particle projected area becomes smaller, and both the forward and backward scattering become smaller (Note that all phase functions in this chapter are normalized in the from of $\frac{1}{2} \int_0^\pi P_{11}(\theta) dcos(\theta) = 1$). The fractal dimension does not show significant influences on other scattering matrix elements, and the differences shown in the P$_{22}$ and P$_{34}$ elements seem noticeable only because of the relatively small ranges used for the y-coordinate. The scattering matrices are much less sensitive to the fractal prefactor, and there are relative differences of less than 20% in the forward and backward directions for phase functions of aggregates with k_f values between 1.0 and 1.8. Thus, the corresponding results are not given here.

As discussed in Sect. 2.3, there are significant uncertainties on the refractive indices of BC particles, and Figs. 13 and 14 illustrate the effects of refractive index on aggregate optical properties. Figure 13 shows the integral scattering properties of aggregates with different refractive indices as functions of aggregate size, and here we use equivalent volume diameter d_v, instead of N, as x-axis. Both real and imaginary parts affect the extinction and absorption of fractal aggregates, and the imaginary part shows a stronger influence than that of the real part. The extinction and absorption show relative differences less than 10% with the real part increasing from 1.6 to 2.0. As the imaginary part increases from 0.4 to 1.0, both the extinction and absorption cross sections are approximately doubled. Thus, the imaginary part should be more carefully studied and constrained for BC optical properties. The bottom panels of Fig. 13 are the asymmetry factors, which are not sensitive to the refractive indices. Because only a single realization is considered for each aggregate size, the asymmetry factors in the bottom panel show obvious oscillations.

Figure 14 illustrates the bulk scattering matrix elements of aggregates with different refractive indices, and results based on three sets of refractive indices are given. Refractive index has almost no effects on the scattering matrix elements. Differences can only be noticed for the ratio of P$_{22}$ to P$_{11}$, whereas all values are close to 1 with relative differences less than 1%.
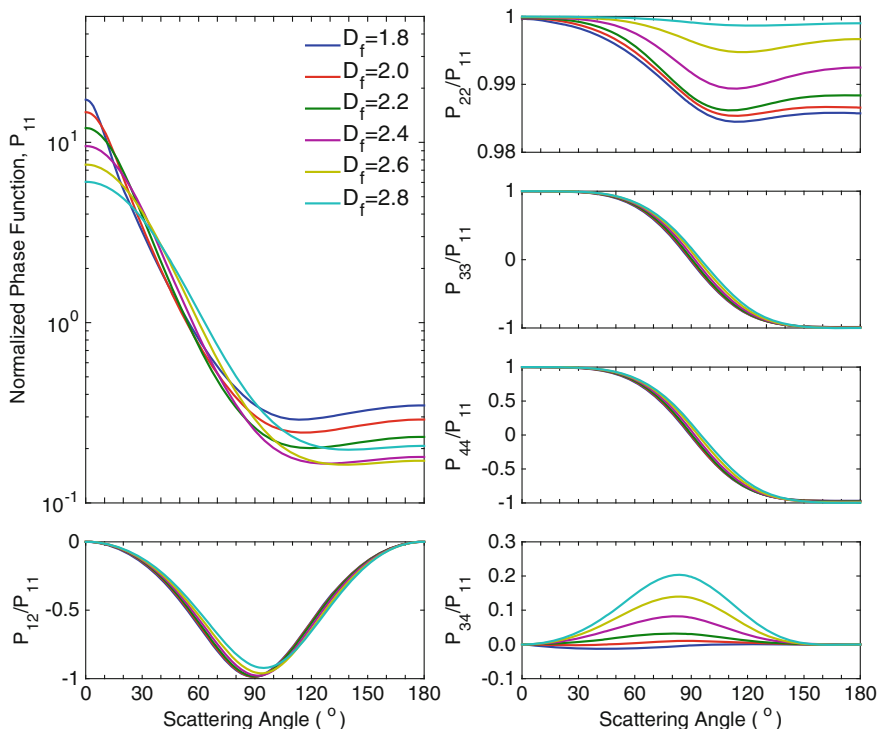
Fig. 12 Bulk scattering matrix elements of aggregates with different fractal dimensions and the default size distributions

This subsection considers the particle overall geometry and refractive index by assuming BC as aggregates of perfectly spherical monomers without overlapping. Besides aggregate size, the fractal dimension and the imaginary part of refractive index show the strongest influences on aggregate optical properties. Meanwhile, different optical properties show different sensitivities on BC geometric parameters or refractive indices. For example, the absorption per volume BC is less sensitive to geometric variables, and the scattering pattern is less sensitive to particle refractive indices.

4.2 Influence of Minor Structures

Different from the perfect FA model, realistic BC monomers are neither same-sized nor attached without overlapping. Some detailed structures do not influence BC overall geometry, whereas may affect BC optical properties to some degrees (Cai et al. 1993; Brasil et al. 1999, 2000; Bescond et al. 2014). How significantly may
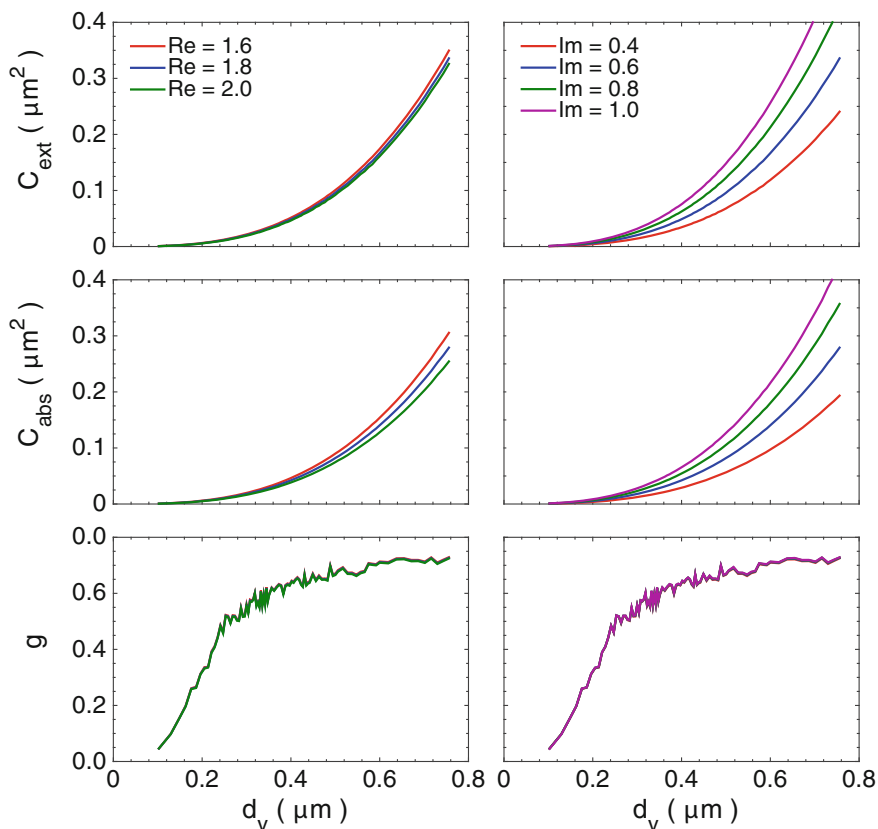
Fig. 13 Effects of refractive indices on the optical properties of black carbon as a function of particle equivalent-volume diameter. The left panels are results with different real parts (Re) but the same imaginary part of 0.6, and the right panels are those with different imaginary parts (Im) but the same real part of 1.8

those structure change BC absorption and scattering properties, and how can they be considered in practical applications? This subsection answers these questions by considering three minor factors as examples: monomer size dispersion, overlapping, and necking. Actually, those features have been investigated in previous studies independently (Farias et al. 1996a; Charalampopoulos and Shu 2002; Bescond et al. 2013; Cheng et al. 2014; Dong et al. 2015; Yon et al. 2015; Doner and Liu 2017), and they obtain quite different conclusions on the influences.

Figure 15 shows examples of aggregates with the minor imperfect structures. The upper panels are aggregates with (a) different-sized monomers, (b) overlapped monomers, and (c) necking between the attached monomers, and the bottom panels are illustrative and detailed sketches to define those structures. To describe monomer dispersion, a size distribution is specified for monomer diameters, such as the red
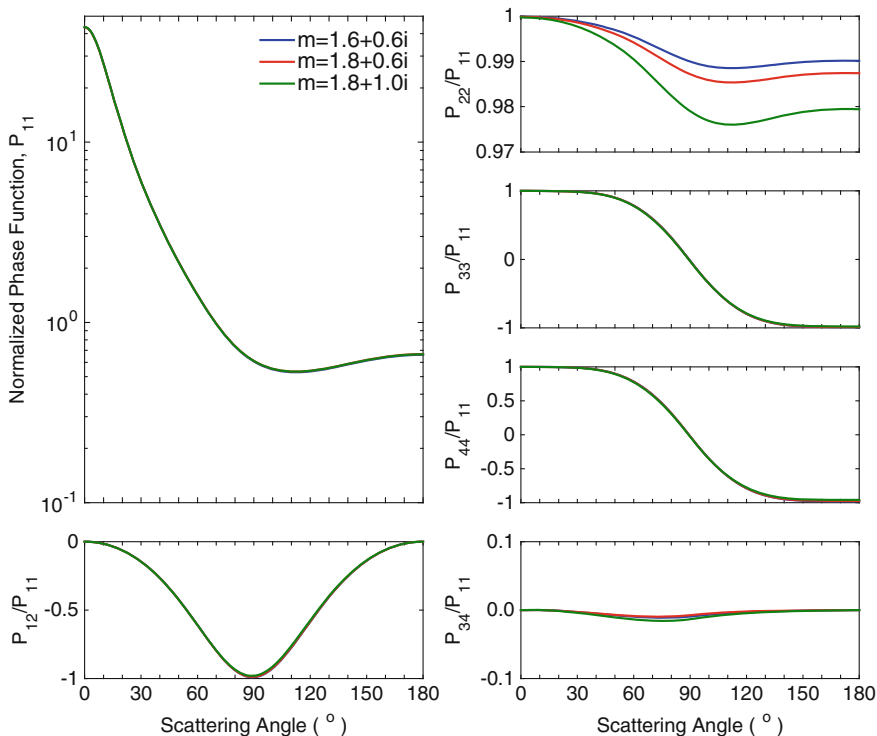
Fig. 14 Non-zero scattering matrix elements of aggregates with different refractive indices. The fractal dimension and prefactor are 1.8 and 1.2 respectively, and monomer diameter is set to be 30 nm

curve in Fig. 3. Then, the diameter of each monomer in an aggregate can be randomly generated to follow the given size distribution, and be aggregated similarly to the FA with same-sized monomers by following the statistic scaling role (Liu et al. 2015a). The monomer overlapping is generated by directly enlarging monomers. In this way, there will be overlapping between monomers, whereas the overall aggregate structure is not effected. The ratio of the enlarged new radius a' to the original one a_o, i.e., a'/a_o, is used to quantify the degree of overlapping (see Fig. 15e). For perfectly attached aggregates without overlapping, we have $a' = a_o$. Note that the monomer themselves are modified, so the effects on the optical properties are not only from overlapping. Thus, the results should be interpreted more carefully. Third, to present the necking, we introduce circular cylinders with the rotational axis coincident to the axis between the two attached monomers, and, as illustrated by the red portion of Fig. 15f, the extra part introduced by the cylinder is referred to as the necking. The radius of the circular cylinder a_n is used to describe the degree of necking, and a_n can vary between 0 and a_o, with $a_n = 0$ representing non-necking. For aggregates with different-sized monomers, spherical monomers are still perfectly attached with each
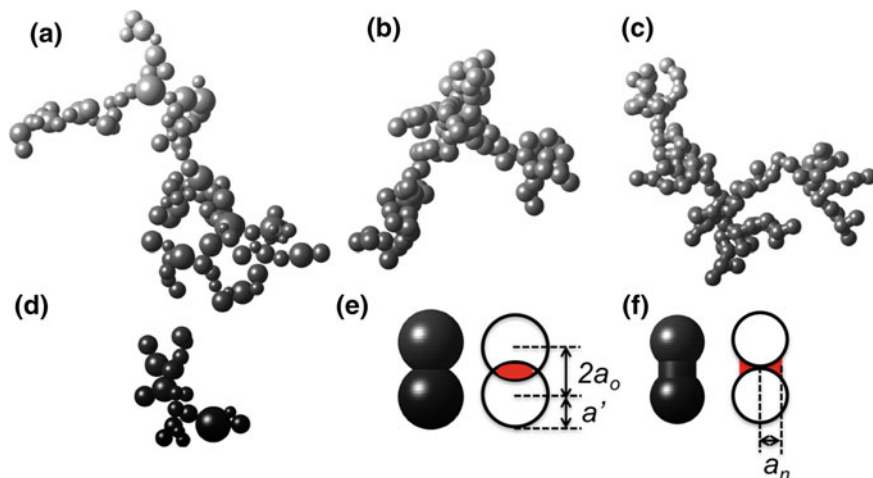
Fig. 15 Examples of fractal aggregates with minor imperfect geometries: **a** different-sized monomers (the lognormal distribution with $d_{\text{geo}} = 33$ nm and $\sigma = 1.25$), **b** overlapping between connected monomers, and **c** necking structures between connected monomers. The fractal parameters of the aggregates are $D_f = 1.8$ and $k_f = 1.2$. The upper panels are for aggregates with 100 monomers, and the bottom panels are illustrative figure to details the minor structures

other, so the MSTM is capable to calculate their optical properties, and the DDA is used for aggregates with overlapping or necking structures.

As mentioned in Sect. 2.2, the lognormal distributions are found to closely represent the observed monomer diameter distributions (Köylü and Faeth 1992, 1994; Lehre et al. 2003; Dankers and Leipertz 2004; Chakrabarty et al. 2006, 2007; Bescond et al. 2014). Figure 16 shows the probability distribution functions of monomer diameter with a GMD of 33 nm and three GSDs (σ) of 1.1, 1.3, and 1.5. As the GSD increases, the monomer size distributions become wider, with not only more larger monomers but also more smaller ones. For $\sigma = 1.5$, the monomer diameters generally range between 10 nm to over 60 nm, and such wide variations of diameter must significantly influence the corresponding optical properties. Actually, σ values from 1.3 up to 1.7 are observed from different fuel combustions (Chakrabarty et al. 2006). More larger monomers appear as the GSD increases, so the aggregate volume and monomer equivalent volume diameter increase. d_v is shown in the right panel of Fig. 16, and the d_v is related to the size distribution by: $d_v = d_{gmd}\sqrt[3]{\exp\left(4.5 ln^2 \sigma\right)}$. With σ increasing from 1.0 to 1.5, the total volume is doubled, and the equivalent volume diameter increases from 33 to ~42 nm.

Figure 17 illustrates the integral optical properties of aggregates with different-sized monomers as functions of the GSD. The blue curves and shaded areas indicate the averaged values and their variation (plus and minus one standard deviation) of the optical properties. A refractive index of $1.75 + 0.435i$ and an incident wavelength of 650 nm are used for this figure based on Liu et al. (2015a). Here, 100 aggre-
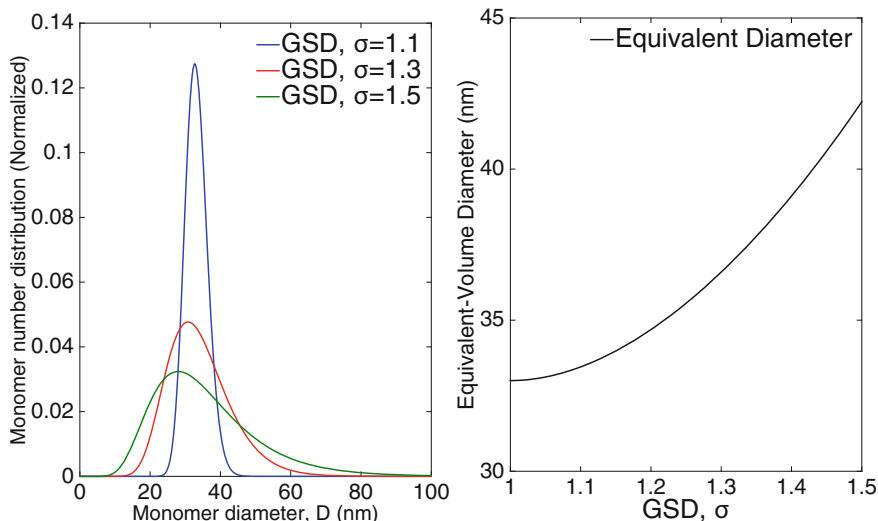
Fig. 16 Left: Lognormal probability distribution of monomer size with geometric mean of 33 nm, and geometric standard deviations of 1.1, 1.3, and 1.5. Right: The equivalent volume diameter of different-sized monomers as a function of geometric standard deviation

gate realizations are generated for each case to account for the randomness of not only aggregation realizations but also monomer size dispersion. As σ increases, the extinction, scattering, and absorption cross sections all increase mainly because of larger total BC volume contributed by larger monomers. The scattering and absorption cross sections at $\sigma = 1.5$ become 2.5 and 1.0 times larger than the corresponding values at $\sigma = 1.0$, i.e., aggregates with same-sized monomers of $2a_o = 33$ nm. The asymmetry factor does not show too much variation over the monomer size deviation. The results in red are the scaled counterparts that remove the effects of changes on monomer equivalent volume diameter by multiplying the absolute value by a correction factor. Considering that the scattering and absorption of a single monomer are proportional to the volume and its square, respectively, the corresponding factors are defined as $(d_v/d_{gmd})^3 = \exp(4.5 ln^2 \sigma)$ and $(d_v/d_{gmd})^6 = \exp(9 ln^2 \sigma)$. For the extinction, we use the same factor as that for absorption, because BC extinction is mainly contributed by its absorption. The changes on the scaled values are much smaller. As σ increases to 1.5, the absorption and scattering after scaling decreases slightly by 5% and 20%, respectively, and there is only a relative variation of less than 0.3% for the extinction. This means that the changes on the absolute cross sections are mainly caused by the increase of the total BC volume, and, thus, can be more easily considered by just multiplying a volume scale factor to the properties of same-sized aggregates. The effects of monomer size distribution on the scattering phase matrix are also relatively weak, and will not be discussed here. More details on the influence of monomer size distribution on BC optical properties can be found in Liu and Smallwood (2011), Wu et al. (2014), and Liu et al. (2015a).
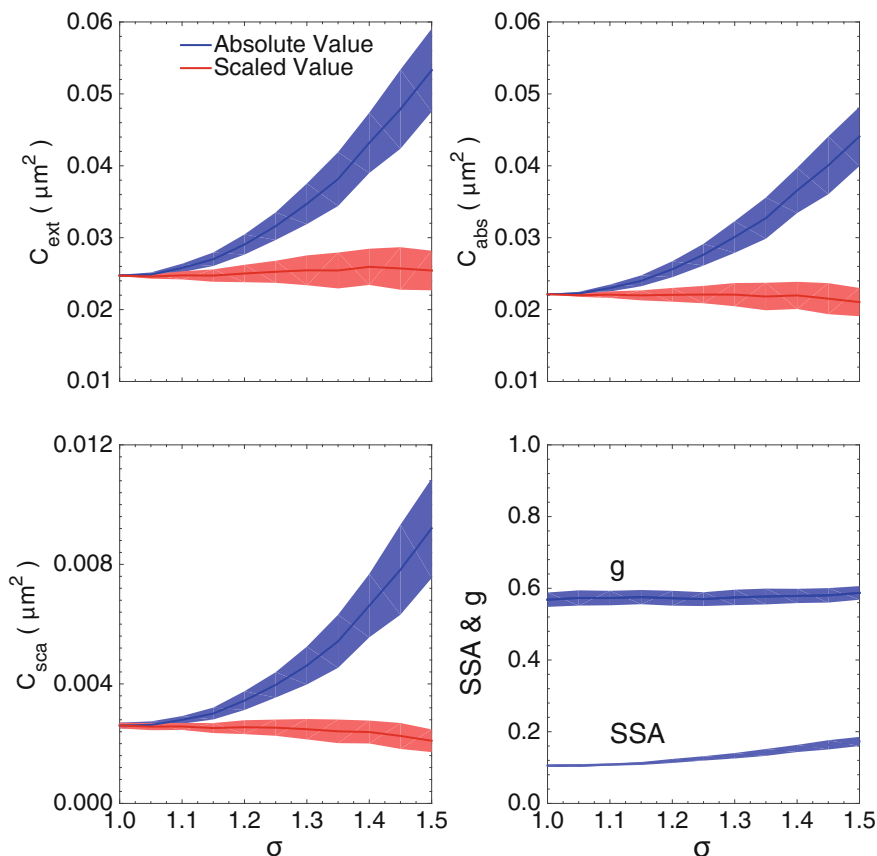
Fig. 17 Optical properties of aggregates with different-sized monomers as functions of geometric standard deviations (σ) of monomer size distributions. Both absolute values (blue) and scaled values (red) (Scale: normalized to account for the change in particle total volume due to standard deviation change). The shaded areas depict the averaged values minus and plus one standard deviation due to aggregate realization

The influences of overlapping and necking on aggregate optical properties are illustrated in Fig. 18. The DDA is used to calculate the optical properties of those 'irregular' particles, and, again, we include both absolute and scaled values in the figure. For the overlapping, the total volume also increases as the overlapping degree increases, i.e., a'/a_o increases. Because the monomers are enlarged, the increases of the absolute cross sections cannot be understood to be contributed only by overlapping. Here, we pay more attestation on the scaled values, which do not change significantly. For the severe necking case with a_n/a_o increasing to 1, the necking volume researches to over 40% of the original aggregate volume, and the extinction and absorption increase over 60% compared to the ideal no-necking case. The scattering cross section shows the most significant variation due to necking, which
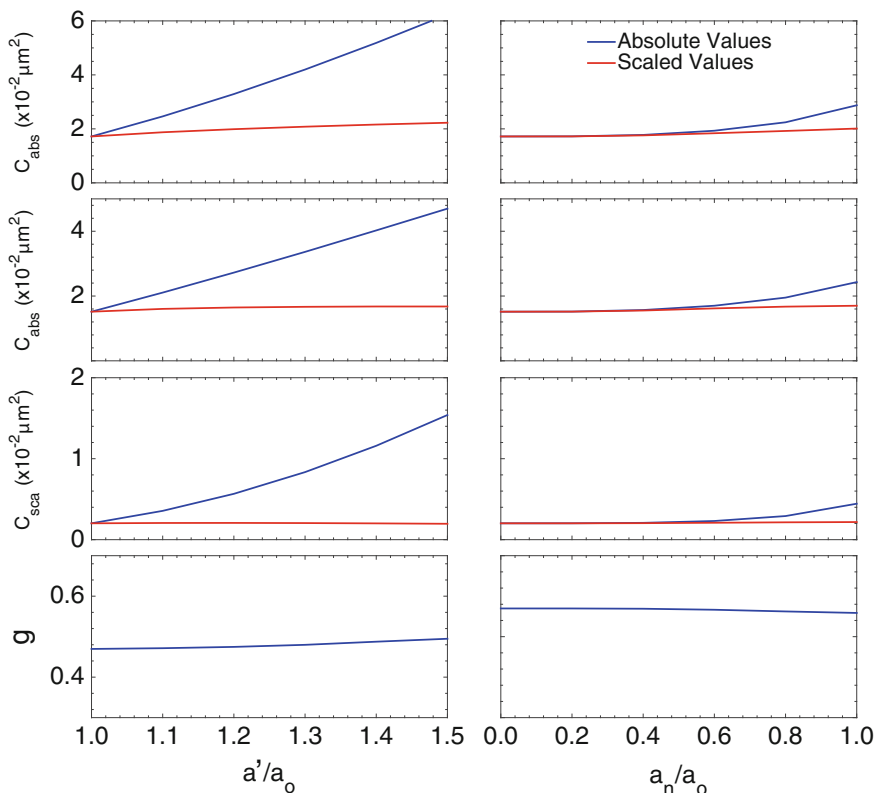
Fig. 18 Optical properties of aggregates with overlapping (left) and necking (right) structures. Both absolute and scaled (accounting for the changes on aggregate total volume) values are given. a_o is original monomer radius, and parameters a' and a_n are used to describe monomer overlapping and necking (see Fig. 15), respectively

is doubled, whereas the asymmetry factor is almost unchanged with the increase of necking. Again, the scaled values (red curves) are those normalized ones considering BC volume, removing effects due to volume change. The influences of overlapping and necking on the scaled extinction and absorption are much weaker, less than 20% for most cross sections. Thus, similar to monomer size dispersion, the effects of overlapping and necking can be roughly approximated by multiplying only a scale factor accounting for the volume differences.

Figure 19 illustrates the scattering matrix elements of aggregates with overlapping $(a'/a_o = 1.1$ and $1.5)$ and necking $(a_n/a_o = 0.2$ and $a_n/a_o = 0.8)$ monomers as well as those of perfect aggregates $(a'/a_o = 1.0)$, and two cases with weak and severe minor structures are illustrated. Again, almost no difference is noticed for the matrix elements besides the phase functions. There are noticeable differences on the phase function and P_{22}/P_{11} of aggregates with heavily overlapping and necking
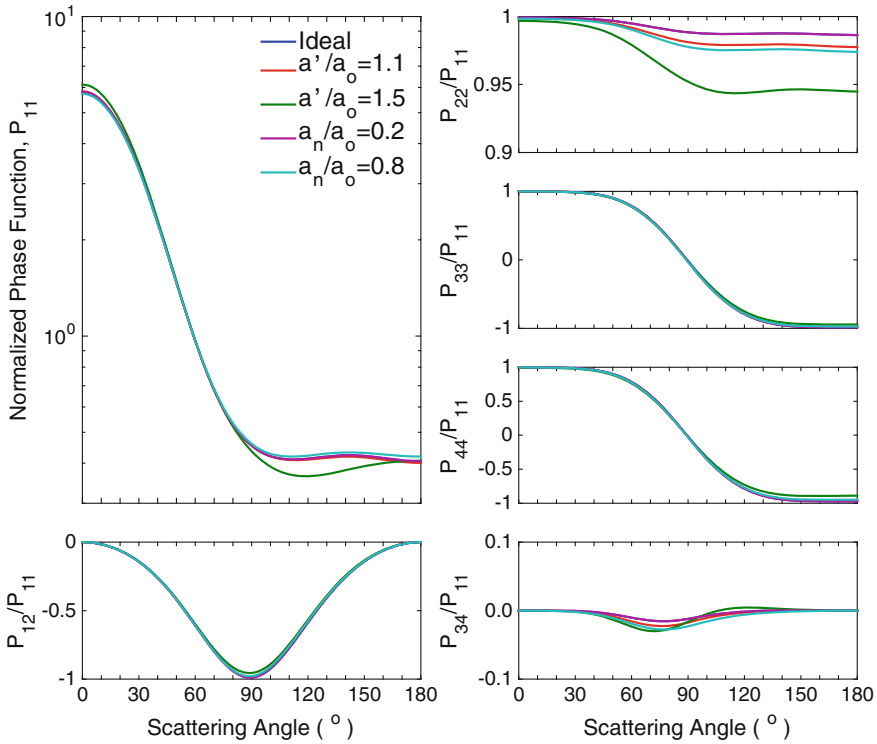
Fig. 19 Comparison of the scattering matrix elements for aggregates with overlapping or necking structures. The aggregate with $N = 100$, $D_f = 1.8$, $k_f = 1.2$ and $a = 30$ is used, and results are those averaged over five aggregate realizations

cases, because, with $a'/a_o = 1.5$ or $a_n/a_o = 0.8$, the aggregates are clearly different from the corresponding perfect case. Considering that we are discussing the effects of minor structures, and imperfect factors with small values of a'/a_o or large values of a_n/a_o are not minor anymore. Note that the P_{22}/P_{11} term may change by ~3% for the severe overlapping case. However, overall, the effects of minor structures on the scattering matrix elements are all minor.

To summarize, the minor imperfect structures do show clear influence on the optical cross sections of BC aggregates, whereas have less effects on the scattering matrix elements, which agree with previous studies. However, the influence is mainly caused by the variation on BC total volume, and scaled results in Figs. 17 and 18 indicate that the optical properties of aggregates with minor imperfect structures can be approximated using those of perfect aggregates by multiplying a correction factor to account for the volume/mass difference.
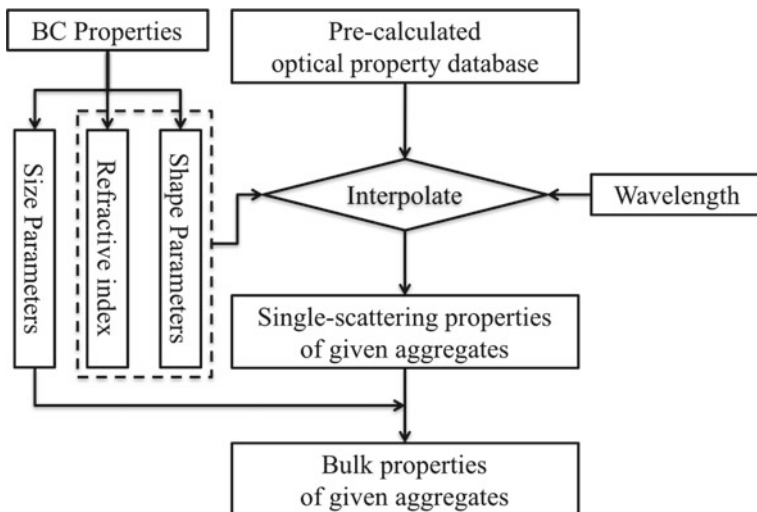
Fig. 20 Framework of the database based on pre-calculated database and interpolation for optical properties of aggregates with small monomers

4.3 An Optical Property Database of BC Aggregates

As we have shown in Sect. 3, either simple RDG approximation or equivalent spherical model introduces obvious errors on aggregate optical properties. Meanwhile, the simulations for optical property of non-spherical particles are computational expensive and trivial, because the aggregates should be rigorously generated and defined. Thus, we extend our simulations for a wide range of aggregate parameters, and develop a database to give optical properties of BC aggregates efficiently and accurately. Figure 20 illustrates the framework of the database. The fundamental role of the database is to obtained BC optical properties with given FA structure parameters, refractive index, and size distribution, and this is done mainly by interpolation among those pre-calculated results given by precise aggregate geometries and the MSTM. The effects of minor imperfect structures are considered by scaling with respect to BC volume as discussed in previous subsection, while a volume correction factor should be approximated and given by the user. Overall, the database includes two parts: a dataset with a large amount of aggregate optical properties precalculated using the MSTM, and an implementation including mainly interpolation to give the properties at certain BC parameters. The bulk optical properties of a given size distribution is given by integral over given size distribution of N or d_v.

Table 7 lists the variables we considered for the large dataset. First, three different fractal dimensions are used to account for both lacy and compact aggregates, while we fix the fractal prefactor to be 1.2 due to its less significant influence on optical properties and relatively less variation. We consider 135 aggregate overall sizes defined by monomer numbers from 1 to 3000. For monomer size, monomer size

Table 7 Variables and parameters considered for pre-calculated optical properties

Parameter	Values
D_f	1.8, 2.3, and 2.8
aggregates size (N)	135 different N values from 1 up to 3000
Monomer size (x_a)	0.05 to 0.5 in steps of 0.05
Real part of refractive index	1.1, 1.2, 1.4, 1.6, 1.8, 2.0
Imaginary part of refractive index	0.1, 0.2, 0.4, 0.6, 0.8, 1.0

parameters (size relative to wavelength), instead of absolute sizes, are set from 0.05 to 0.5 in steps of 0.05, so different-sized monomers that have the same size parameters at different wavelengths are calculated only once in the database. Six real parts and six imaginary parts of refractive indices are used, and this leads to a total of 36 complex refractive indices. Thus, the optical properties of a total of almost 150,000 aggregates with different parameters are calculated by the MSTM. The details of the database can be found in Liu et al. (2019), and much more results are discussed in the study.

Figure 21 illustrates an example of the pre-calculated optical properties. The top panels are the optical properties of aggregates with a fractal dimension of 1.8 and refractive index of $1.8+0.4i$. The x-axis and y-axis are for aggregate size N and monomer size parameter x_m, respectively. Because the area of corresponding equivalent volume sphere, instead of the project area of the aggregate, is used to calculate the extinction efficiency, the efficiencies reach as large as 8 in the figure. A factor of 10 is multiplied to the *SSA* and *g* values in the figure, so the same color bar as those for Q_{ext} with values between 0 and 8 can be used. The middle and bottom panels are those of more compact aggregates with D_f of 2.3 and 2.8, respectively, and the results are given as their relative differences from those of aggregates with $D_f = 1.8$ for comparison. All optical properties increase as the monomer size or monomer number increases, because most of those aggregates have equivalent volume size parameters less than 2. The middle and bottom panels show the influence of fractal dimension on the optical properties, and the influences are substantially different, which will not be discussed in details. We calculate similar data for all 36 different refractive indices, and that makes it possible to give optical properties for most BC related applications.

5 Optical Properties of Mixed BC

As mentioned in Sect. 2.1, after emitted into the atmosphere, BC particles may mix with other aerosols, and become inhomogeneous with coatings of nitrate, sulfate, or other carbonaceous materials, most of which are non-absorbing. This process is also known as aging, and the mixed BC can also be referred to as aged BC. The
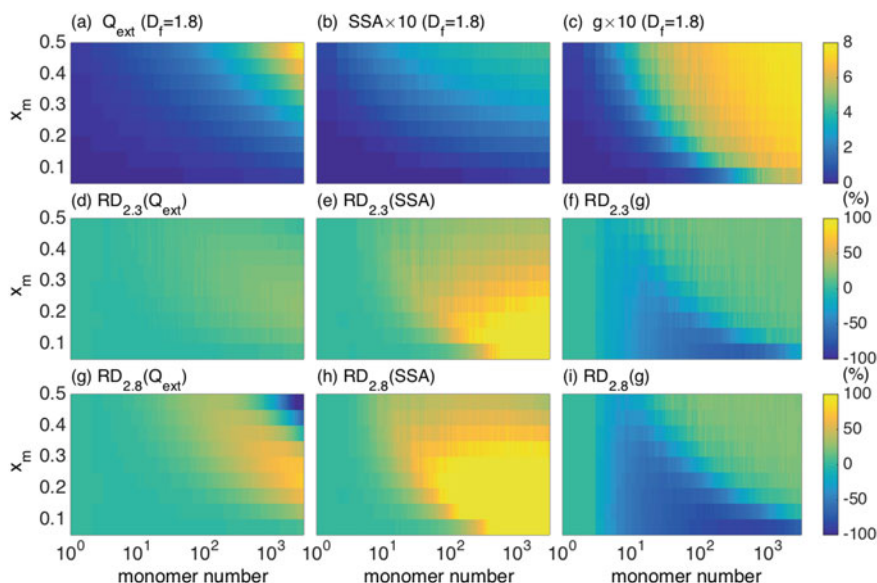
Fig. 21 Top panel: Extinction efficiency, single-scattering albedo, and asymmetry factor of fractal aggregates with a fractal dimension of 1.8 and a refractive index of $1.8 + 0.4i$ at different monomer and aggregates sizes. Middle and bottom panels: Relative differences (RD) of the scattering properties for aggregates with different fractal dimensions

mixing state brings another important, if not the largest, uncertainty on BC optical properties. This section answers two questions: how the mixing state and mixing properties would affect BC optical properties, and how significant non-absorbing coating may enhance BC absorption.

The mixing states, e.g., size, structure, component, and relative position, make the BC particles even more complex. For simplification, this study considers only non-absorbing coating. If brown carbon, which is moderately absorbing itself, is coated on strongly absorbing BC, its influence on the optical properties, especially total absorption, may be even more difficult to be evaluated. Due to the uncertainties on brown carbon or other absorbing materials, we consider only non-absorbing coating in this chapter, e.g., sulfate aerosols, and a corresponding refractive index of 1.45 at the wavelength of $0.5 \, \mu m$ (imaginary part of zero) is considered.

5.1 Internally Mixing Model

To account for the effects of BC aging, an ideal numerical model is needed to represent the internally mixed BC aggregates. As we have shown in Table 4, the DDA takes much more computational time for the simulations with mixed BC particles, so we

want to take the advantage of the MSTM efficiency. This is possible, because the MSTM can consider clusters with spheres that are totally inside another sphere. In the model, we keep using the concept of the FA for BC particles, and simplify the coating to have a spherical overall geometry following the model developed by Liu et al. (2017a). For heavily coated BC, the coating does have spherical coating (see Fig. 1n, o). Even with this simplest case, there are at least three more parameters to specify the particle, i.e., the component (fixed to be sulfate), size, and relative position of the coating sphere and aggregate. Three steps are used to generate a coated particle. First, a bare BC aggregate and a sphere with given sizes are generated, and the size/volume of the coating sphere is calculated by given BC volume and its volume fraction in the mixture. Second, the coating sphere and BC aggregate are mixed as a superposition with a defined relative distance between the two particles. Third, the coating sphere and aggregate are adjusted for the sake of applying the MSTM. Because some volume of the coating sphere is occupied by BC monomers inside the sphere, its radius should be enlarged until researching the required coating volume. Meanwhile, to avoid the overlapping of the BC monomer and coating sphere, we randomly remove all overlapped monomers (partially inside and partially outside) to the outside of the coating sphere, and the relocation is performed to make the movement as short as possible. In this way, the overall structure is not significantly modified, and the generated particle can be considered by the MSTM.

Figure 22 shows some examples of modeled FAs with spherical coating, and both lacy and compact aggregates are illustrated in the figure. The large transparent spheres in the figure represent the coating material, and the BC aggregates all have 100 monomers. In the top and middle panels, the volume fraction of the coating sphere from left to right increases from 0.5 to 0.9, and the centers of the coating sphere are located at mass center of the aggregate. The bottom panels show 'coating' attached at the edge of a BC aggregate, and they can be understood as an extreme mixing case. In the MSTM, those coated aggregates are treated as a single inhomogeneous particle, and the interaction between its two components are fully accounted. This section adopts this special model to represent mixed BC aggregates, and, considering the significant variations of coating geometries that may occur in the atmosphere, multiple models should be combined to account for more realistic effects of coating.

5.2   Influence of Coating Properties

The influence of aging/coating on BC optical properties has become one of the most popular topics on BC radiative forcing and retrieval, because mixing properties as well as the corresponding optical properties are still poorly understood (Moffet and Prather 2009; Wu et al. 2016; Liu et al. 2013, 2017b). As a result, it is necessary to use idealized models to capture some important factors for further investigation and to understand the effects quantitatively. The simple model introduced above is considered, because the corresponding particles are similar to realistic particles and their optical properties can be efficiently and accurately calculated by the MSTM.
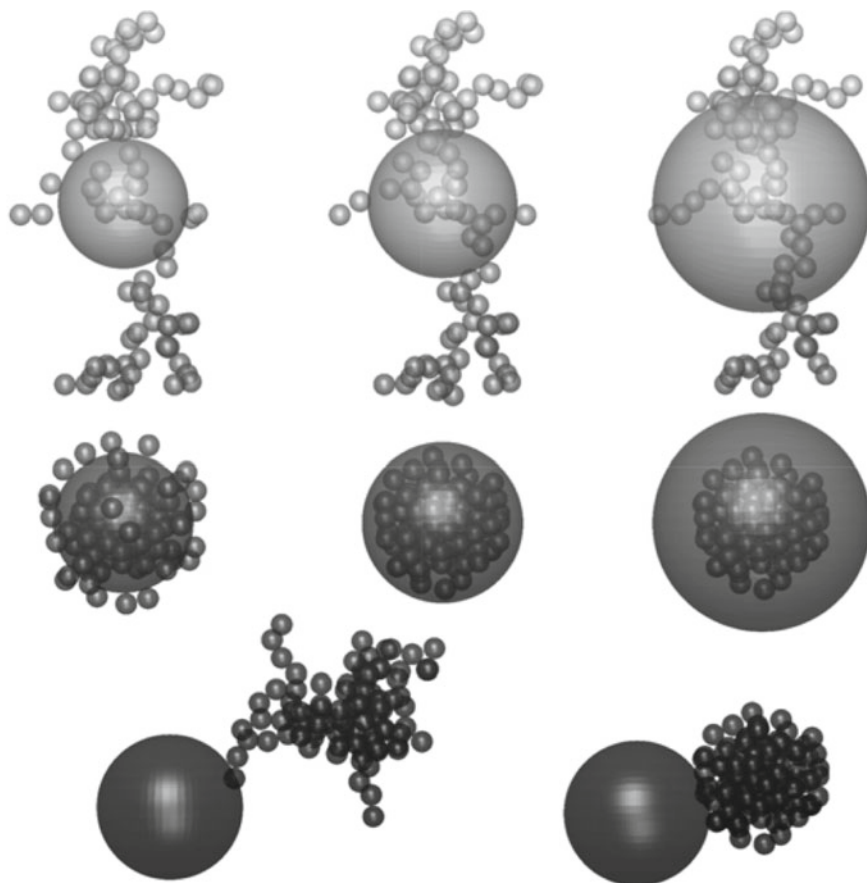
Fig. 22 Examples of aggregates with spherical coating. Top and middle panels: the coating volume fraction changes from 0.5, 0.75, to 0.9 (from left to right). Bottom panels: aged BC with the coating sphere attached at aggregate edge

Figure 23 illustrates the effects of coating on the integral optical properties, and we consider (1) coating volume fraction F_{coating} (left) and (2) relative position between the centers of BC aggregate and coating sphere (right). BC aggregates with 100 monomers and D_f of 1.8 (Lacy) and 2.8 (Compact) are considered. In the left panels, we fix the size and shape of the BC aggregates, and increase the coating volume (coating volume fraction, i.e., F_{coating}). The extinction and scattering cross sections show the most dramatic variations, because the coating increases particle sizes. More importantly, the absorption of lacy and compact aggregate increases by 50% and 100% respectively, which is widely known as the "lensing effects" (Fuller et al. 1999; Schnaiter et al. 2005; Peng et al. 2016), and the enhancement is more significant as the coating fraction becomes larger than 0.5. The degree of enhancement can be explained by how significantly the BC aggregates are embedded by the coating. For

small coating fraction, only a small fraction of the aggregate is embedded. We will discuss this enhancement on absorption with more details in the following subsection. With the increase of the non-absorbing material, the single scattering albedo of the mixture increases from approximately 0.1 to over 0.8, and an absorbing dominant particle becomes scattering dominant. The changes on the asymmetry factor are also reasonable as the particle becomes larger with the increase of the coating fraction (again, BC particle is fixed). In the right panels, we fix coating volume fraction at $F_{coating} = 0.9$, and change the relative distance between the two. The x-axis is for normalized distance between BC aggregate and coating. A normalized distance of 0 represents a coating sphere located at the aggregate mass center, and 1 corresponds to the extreme case with coating sphere attached at the 'edge' of the aggregate (e.g., the bottom panels of Fig. 22). With the coating moving to the edge of the aggregate, the extinction decreases by approximately 15%, and the absorption of compact and lacy aggregates decreases by 50% and 25%, respectively. However, the effects of the coating relative position on scattering cross section and asymmetry factor are less significant.

Figure 24 shows the scattering matrix elements of a bare aggregate, a coated one, and the corresponding coating sphere. The aggregate has 100 monomers with a radius of 15 nm, and the volume fraction of the coating is 0.9, which makes the coating sphere to be the dominate component of the particle. The scattering patterns are mainly determined by those of the large coating sphere, especially for the forward scattering, and the coated aggregate gives larger backward scattering. The phase functions of 'coated lacy' and 'coated compact' aggregates are close to each other, so we only show those of lacy ones. Other scattering matrix elements show similar properties. Thus, the scattering matrix elements of mixed BC aggregate are determined by those of both the aggregate and the coating sphere, and the relative volume of the two components become an important parameter.

Although this simple model is helpful to understand the effects of coating on BC optical properties, the uncertainties on aged BC make the interpretation difficult. To be more specific, without clear knowledge on the detailed mixing states, numerical model itself can hardly quantify the coating effects.

5.3 How Significantly May Coating Enhance BC Absorption

The influence of coatings on BC absorption has been extensively studied both theoretically and experimentally (Fuller et al. 1999; Schnaiter et al. 2005; Shiraiwa et al. 2010; Dong et al. 2015; Peng et al. 2016; Liu et al. 2013, 2017b). The coating is well-known to enhance absorption of BC, whereas observations with different instruments from different circumstances obtain different results with a factor ranging from 1.3 to 4 times (Schnaiter et al. 2005; Bond et al. 2006; Cappa et al. 2012). This enhancement may lead to systematic errors on understanding BC properties, e.g., measurements of BC concentrations, because some instruments infer BC properties based on absorption measurements. Different optical models for the internal
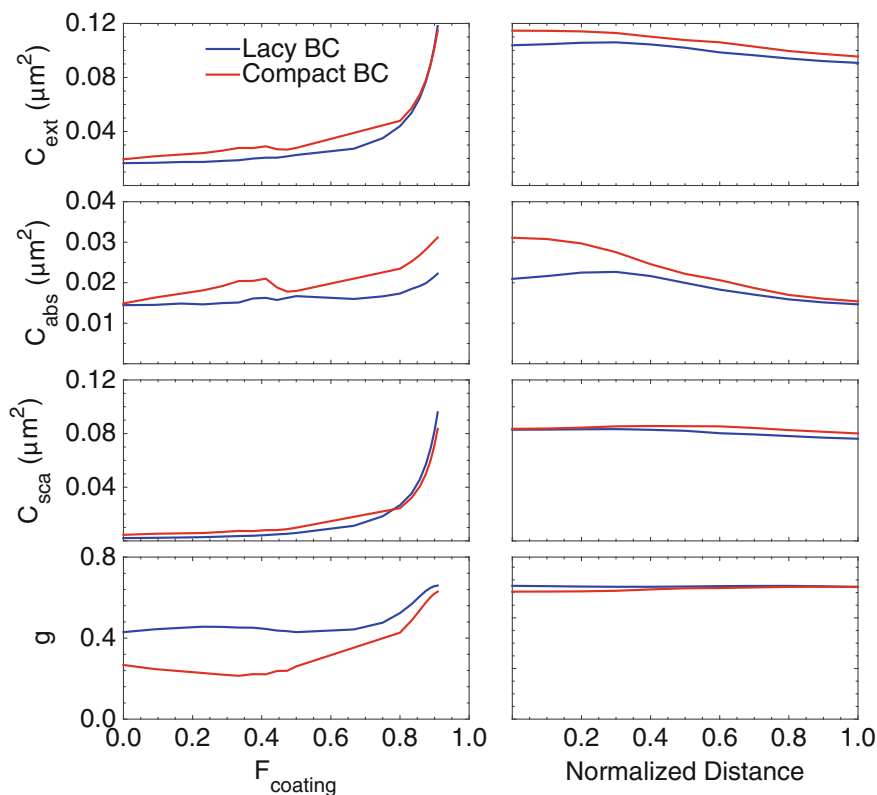
Fig. 23 Optical properties of black carbon aggregates with 100 monomers as functions of coating volume fraction $F_{coating}$ (left panels) and normalized distance between BC and coating centers (right panels)

mixture between BC and non-absorbing component have also been applied to tackle the problem related to absorption enhancement (Lesins et al. 2002; Chung et al. 2012b; Dong et al. 2015), and simple examples include the effective medium theories and concentric core-shell model. However, the experimental measurements of atmospheric BC conducted by Cappa et al. (2012) show that the measured absorption enhancement due to weakly absorbing coating is much lower than that predicted by the core-shell Mie theory. These inconsistent findings and different observations suggest that the effects of coating on BC absorption are still not well understood, and the numerical model may provide some hints.

Previous discussions consider BC aggregates with spherical coating material, which is a reasonable and practical model for aged BC, whereas BC particles in the atmosphere show much larger variations on their geometries. Transmission or scanning microscopic images show almost any geometry you can imagine, and a single model is impossible to represent all possible particles in the ambient atmosphere.
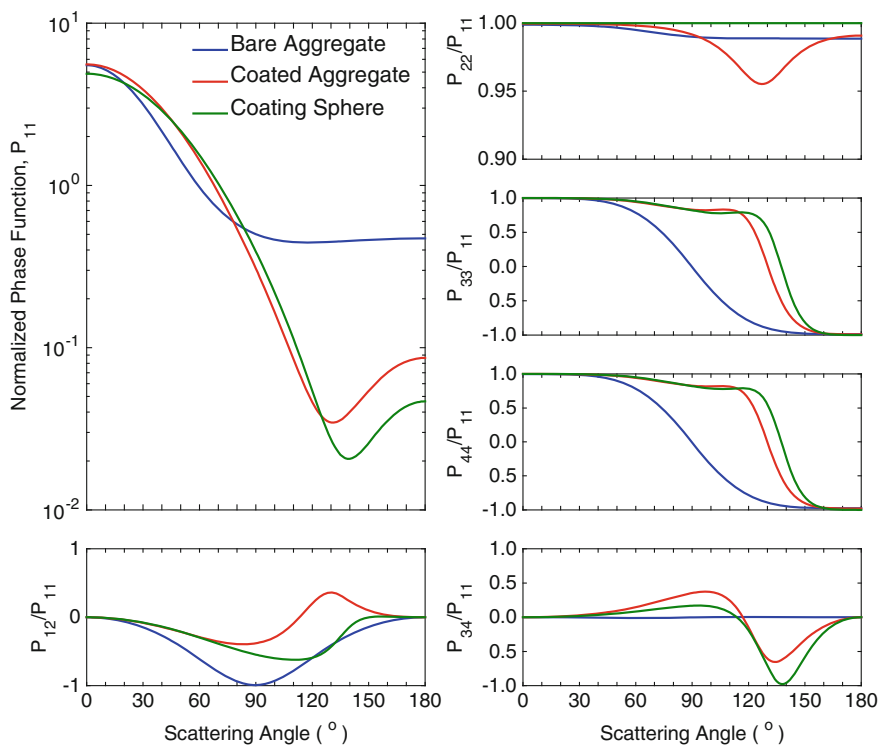
Fig. 24 Scattering matrix elements of a bare aggregate, a coated aggregate, and a coating sphere

Seven inhomogeneous models with different mixing structures are considered here as shown by Fig. 25. Most of those models have been considered on other studies independently to account for the effects of aging or internal mixing on BC optical properties, whereas these results can hardly be compared with each other due to the different definitions on the particle properties. Thus, different models show influences on BC optical properties to different degrees, and an inter-comparison of different models may present a more realistic representation on BC states in the atmosphere. The BC components are treated as either spheres or aggregates, while the coating geometries are considered to be spherical in this study for the sake of applying the MSTM.

Figure 26 shows the averaged absorption cross sections of the bare and mixed BC particles, and the coating amount is assumed to give a coating volume fraction of 0.9. The two columns in the left are for bare BC with lacy and compact structures, respectively, and the other cases are for aged BC with different mixing structures corresponding to the geometries from Fig. 25c–i. Case 6 gives the largest absorption enhancement, which assumes BC monomer randomly located in the coating sphere. Cases 5 and 8 give similar enhancement, in which BC particles are located in the coating center in the form of compact aggregates and sphere, respectively. Case 3
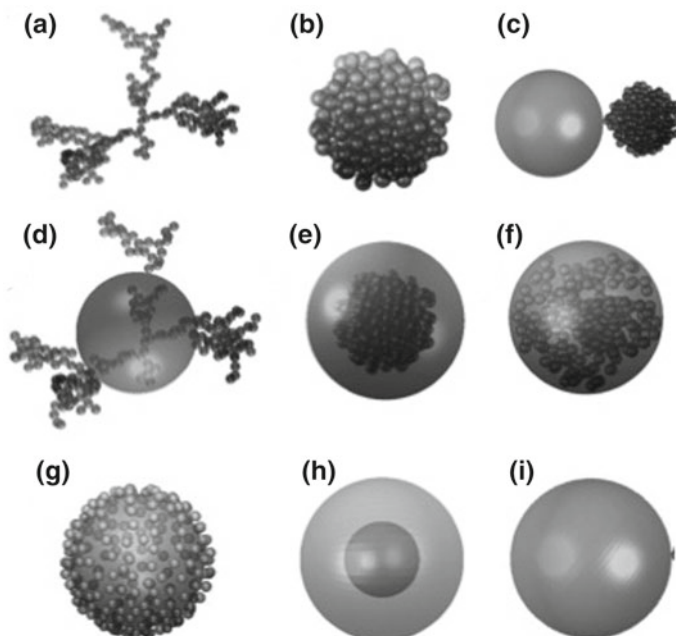
Fig. 25 Particles geometries considered to represent the internal mixing of non-absorbing coating and BC

shows almost no increase in the absorption, because the particles are 'externally' mixed and the interaction between BC and coating is relatively weak. For Case 7, the BC on the coating surface limits the energy reaching the particle, so the absorption is also not significantly enhanced. Overall, as Fig. 26 shows, the mixing with non-absorbing material may enhance the absorption by a factor from 5% to approximately 100% due to different mixing and BC structures. This range of enhancement may also change due to further increase in the coating volume or properties of coating material. For example, hydroscopic aerosols, e.g., sulfate or nitrate aerosols, can easily mix with BC, and mixed BC may absorb water vapor in the humid environment. Thus, an absorption enhancement larger than 100% may be reasonable if BC particles are heavily coated by absorbing a large amount of water. Again, it should be noticed that this study assumes a fixed volume fraction of BC or coating for different-sized BC particles, which may be not true, and this may also influence the results to some degree.
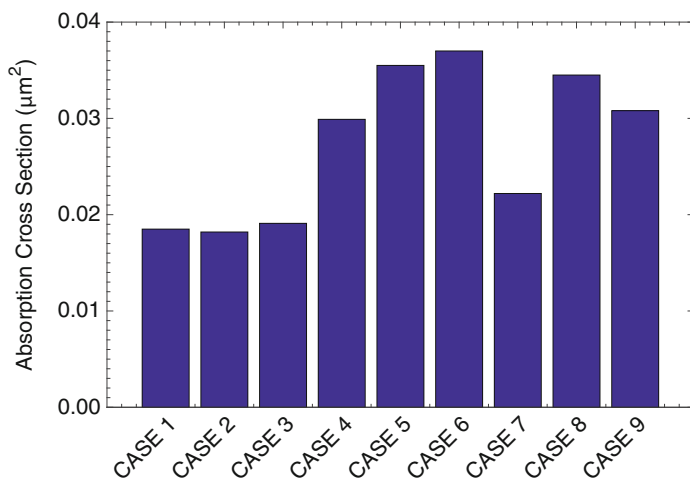
Fig. 26 Averaged BC absorption cross sections for uncoated (Cases 1 and 2) and coated (Cases 3–9) black carbon aggregates with different particle geometries. Cases 1–9 correspond to geometries illustrated in Fig. 24, and volume fraction of coating is 0.9

6 Is BC Black in the Atmosphere?

6.1 Numerical Model

With the discussion of BC, we can hardly ignore the existence of a relatively less absorbing carbonaceous material, i.e., Brown Carbon (BrC), which is named based on its appearance color and spectral absorptivity. BrC and BC are separated by their observational colors, and have quite different spectral absorptivity variation. BC is referred to as 'black' carbon, because it is highly absorbing and shows black color (Bond 2001; Bond et al. 2013; Chakrabarty et al. 2013). Does BC only show black color in the atmosphere? It is difficult to give an accurate answer from observation, because aerosols in the atmosphere are always a mixture of different components. Meanwhile, the BrC that has a stronger spectral variation on its absorption is supposed to show brown color in the atmosphere (Andreae and Gelencsér 2006) due to its stronger absorption in the shorter visible wavelengths, and its absorption decreases more dramatically in the visible regions as wavelength increases (Kirchstetter et al. 2004; Bahadur et al. 2012; Gyawali et al. 2012). Thus, it is interesting to investigate the realistic color of BC and other aerosols in the ambient atmosphere.

We simulate the color of a BC aerosol layer in the ambient atmosphere in three steps. First, we carry out single scattering simulations to obtain the scattering properties of BC particles, and the bulk optical properties are obtained by averaging over typical BC size distributions. Secondly, radiative transfer simulations are performed to approximate the portion of visible light that goes through the aerosol layer

(a)

(b)
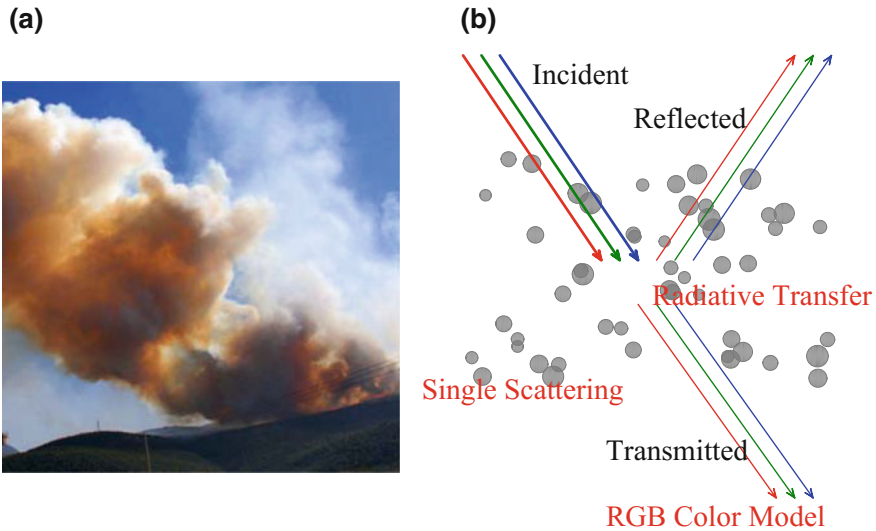
Fig. 27 **a** A colorful smoke image of fire, and **b** configuration to simulate colors of BC layer in the atmosphere. The transmittance (instead of reflectance) of the red, green and blue light is used in this study for the color display, and the models used to display color is mentioned in the figure in red

and reaches the surface at each single wavelength. A two-stream radiative transfer model is used, and it can efficiently approximate the reflectance, transmittance, and absorptance of an aerosol layer. Last, the transmitted light at the visible spectrum is visualized based on the RGB color model to show its color in the ambient atmosphere (Liu et al. 2016a).

The left panel of Fig. 27 is a colorful image of fire smoke, and we can see colors such as yellow, brown, or grey colors from white to black. The right panel of Fig. 27 illustrates the procedure of our simulation. It is similar to the realistic ambient condition, in which solar light reaches the observer after passing an aerosol layer. In the figure, the incident light, i.e. purely white light, will be scattered and absorbed by the aerosol layer, and the transmitted light, including both the direct and diffuse energy, is used to display the color seen by an observer. Random spheres in the figure represent the aerosol layer, and their optical properties determine the relative strength of each outgoing monochromatic light, then the color. This procedure is mentioned with much more details in Liu et al. (2016a) to understand carbonaceous aerosol colors.
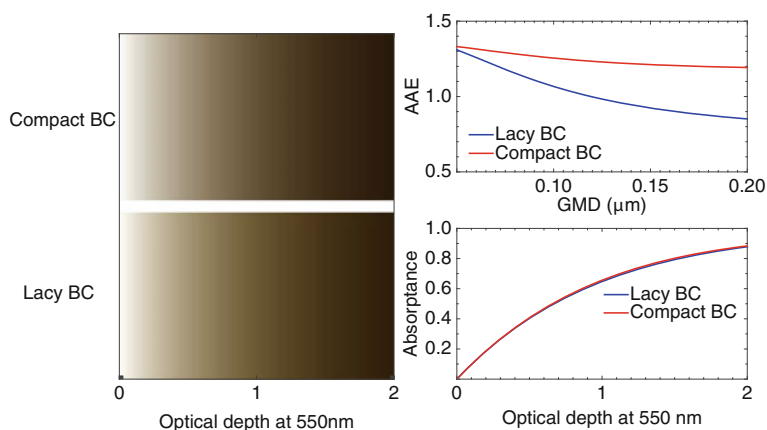
Fig. 28 Left: Colors of black carbon aerosol layers as a function of optical depth (wavelength of 550 nm). Right: Absorption Ångström Exponent (AAE) (top) and absorptance (bottom) of the aerosol layer at the wavelength of 550 nm

6.2 BC Color in the Atmosphere

Figure 28 shows the colors of black carbon aerosols with realistic particle size distributions as a function of optical depth at the wavelength of 550 nm, and the Absorption Ångström Exponent (AAE) (Russell et al. 2010) and absorptance of the BC layer are shown in the right panels. The default aggregate size distribution is used for the optical property simulations. We consider aerosol layers with optical depths at wavelength of 550 nm ranging from 0 up to 2. The incident angle is set to 40°. The wavelength-dependent refractive indices from Chang and Charalampopoulos (1990) are considered for BC optical properties. The AAE is a parameter describing the spectrum variance of aerosol absorption, and larger AAE means that the absorption decreases more quickly as the wavelength increases (Ganguly et al. 2005; Lack and Cappa 2010; Lack and Langridge 2013).

Thus, stronger spectral variation on the absorption leads to stronger differences on the transmitted light and gives brown color for the BC layer. It is expected that, as the optical depth increases, the color of the aerosol layer becomes darker, because the layer reflects and absorbs more and more incident energy. Furthermore, both lacy and compact BC show some brown colors with optical depth under 1, whereas become almost black with optical depth larger than ~1.2. As shown in the right top panel, both lacy and compact BC have AAE larger than 1 at a GMD of 0.12 μm. The black color is shown for optical depth larger than ~1.2, which corresponds to absorptance greater than 0.8. This indicates that the aerosol layer cannot be too thick to show 'colors' other than black.

Figure 29 illustrates the colors of BC aerosols at different particle sizes (i.e., different geometric mean diameters) and optical depths, and the BC particles are assumed to be lacy (left) and compact (right) aggregates, respectively. It clearly
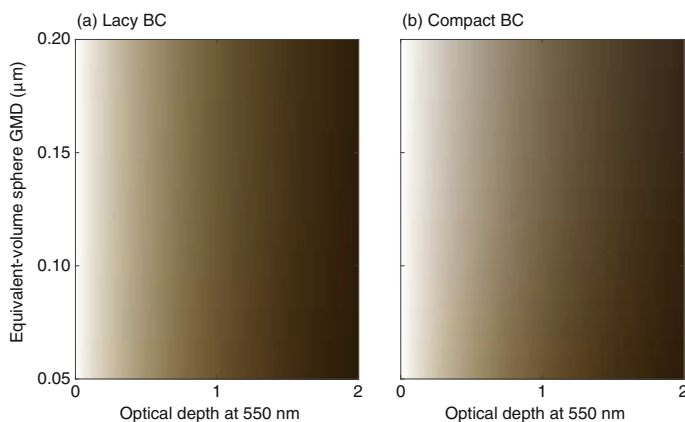
Fig. 29 Colors of lacy (left) and compact (right) black carbon layers at different optical depth and particle geometric mean diameters. The optical depth is for the value of the aerosol layer at the green wavelength 550 nm

demonstrates that the color of aerosol layers is less sensitive to BC geometry or size distribution than to optical depth, whereas brown colors are noticed for all BC sizes. To summarize, although BC is named as black carbon, it may show slight brown colors in the ambient atmosphere.

7 Conclusion

This chapter focuses on the scattering and absorption properties of BC aerosols, and reveals the importance of particle nonsphericity and inhomogeneity. Geometrically, BC particles are represented using the fractal aggregates, and various numerical models (e.g., the RDG, equivalent-sphere approximations, MSTM, GMM, and DDA) with quite different accuracies are capable to simulate their optical properties. Those models used for light scattering properties of BC aggregates are compared, and, by considering both the efficiency and accuracy, the MSTM shows better performance. The geometries of fractal aggregates are also modified to account for effects of particle minor structures (polydisperse monomer, overlapping, and necking) on the optical properties, and a simple numerical model is developed to consider the effects of non-absorbing coating. Our results indicate that the nonsphericity and inhomogeneity of BC particles should be considered for further applications such as BC property retrievals and radiative transfer.

As we have discussed in this chapter, the numerical models show great capability on calculating BC optical properties, and it's possible to treat black carbon particles as almost any ideal geometries (e.g., fractal aggregates, minor structures, internal mixing), which definitely improve the representation of natural BC particles. How-

ever, considering the significant uncertainties related to BC microphysical properties and variations on BC particle themselves, numerical modeling itself can never give us a complete understanding on BC optical properties, and more efforts should be devoted to the collaboration of observation and numerical modeling. Only by considering observations with more details, the numerical model can help to improve our understanding on BC optical properties. Meanwhile, the numerical results should be used more frequently to verify observations.

References

Adachi K, Chung SH, Buseck PR (2010) Shapes of soot aerosol particles and implications for their effects on climate. J Geophys Res 115. https://doi.org/10.1029/2009jd012868

Alexander DT, Crozier PA, Anderson JR (2008) Brown carbon spheres in East Asian outflow and their optical properties. Science 321:833–836

Andreae MO, Gelencser A (2006) Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. Atmos Chem Phys 6:3131–3148

Arnott WP, Hamasha K, Moosmüller H et al (2005) Towards aerosol light-absorption measurements with a 7-wavelength aethalometer: evaluation with a photoacoustic instrument and 3-wavelength nephelometer. Aerosol Sci Tech 39:17–29

Bahadur R, Praveen PS, Xu Y et al (2012) Solar absorption by elemental and brown carbon determined from spectral observations. Proc Natl Acad Sci USA 109:17366–17371

Bambha RP, Michelsen HA (2015) Effects of aggregate morphology and size on laser-induced incandescence and scattering from black carbon (mature soot). J Aerosol Sci 88:159–181

Bergstrom RW, Russell PB, Hignett P (2002) Wavelength dependence of the absorption of black carbon particles: predictions and results from the TARFOX experiment and implications for the aerosol single scattering albedo. J Atmos Sci 59:567–577

Bergstrom RW, Pilewskie P, Schmid B et al (2003) Estimates of the spectral aerosol single scattering albedo and aerosol radiative effects during SAFARI 2000. J Geophys Res 108. https://doi.org/10.1029/2002jd002435

Bescond A, Yon J, Girasole T et al (2013) Numerical investigation of the possibility to determine the primary particle size of fractal aggregates by measuring light depolarization. J Quant Spectrosc Radiat 126:130–139

Bescond A, Yon J, Ouf FX et al (2014) Automated determination of aggregate primary particle size distribution by TEM image analysis: application to soot. Aerosol Sci Tech 48:831–841

Bibi S, Alam K, Chishtie F et al (2017) Observations of black carbon aerosols characteristics over an urban environment: radiative forcing and related implications. Sci Total Environ 603–604:319–329

Bohren CF, Huffman DR (2008) Absorption and scattering of light by small particles. Wiley, New York

Bond TC (2001) Spectral dependence of visible light absorption by carbonaceous particles emitted from coal combustion. Geophys Res Lett 28:4075–4078

Bond TC, Bergstrom RW (2006) Light absorption by carbonaceous particles: an investigative review. Aerosol Sci Tech 40:27–67

Bond TC, Sun H (2005) Can reducing black carbon emissions counteract global warming? Environ Sci Tech 39:5921–5926

Bond TC, Covert DS, Kramlich JC et al (2002) Primary particle emissions from residential coal burning: optical properties and size distribution. J Geophys Res 107. https://doi.org/10.1029/2001jd000571

Bond TC, Habib G, Bergstrom RW (2006) Limitations in the enhancement of visible light absorption due to mixing state. J Geophys Res 111. https://doi.org/10.1029/2006jd007315

Bond TC, Bhardwaj E, Dong R et al (2007) Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850–2000. Glob Biogeochem Cycle 21. https://doi.org/10.1029/2006gb002840

Bond TC, Zarzycki C, Flanner MG et al (2011) Quantifying immediate radiative forcing by black carbon and organic matter with the specific forcing pulse. Atmos Chem Phys 11:1505–1525

Bond TC, Doherty SJ, Fahey DW et al (2013) Bounding the role of black carbon in the climate system: a scientific assessment. J Geophys Res 118:5380–5552

Brasil AM, Farias TL, Carvalho MG (1999) A recipe for image characterization of fractal-like aggregates. J Aerosol Sci 30:1379–1389

Brasil AM, Farias TL, Carvalho MG (2000) Evaluation of the fractal properties of cluster-cluster aggregates. Aerosol Sci Tech 33:440–454

Burr DW, Daun KJ, Thomson KA et al (2012) Optimization of measurement angles for soot aggregate sizing by elastic light scattering, through design-of-experiment theory. J Quant Spectrosc Radiat 113:355–365

Cai J, Lu N, Sorensen CM (1993) Comparison of size and morphology of soot aggregates as determined by light scattering and electron microscope analysis. Langmuir 9:2861–2867

Cappa CD, Lack DA, Covert DS et al (2008) Bias in Filter-based aerosol light absorption measurements due to organic aerosol loading: evidence from ambient measurements. Aerosol Sci Tech 42:1033–1041

Cappa CD, Onasch TB, Massoli P et al (2012) Radiative absorption enhancements due to the mixing state of atmospheric black carbon. Science 337:1078–1081

Chakrabarty RK, Moosmüller H, Garro MA et al (2006) Emissions from the laboratory combustion of wildland fuels: particle morphology and size. J Geophys Res 111:1135–1153. https://doi.org/10.1029/2005JD006659

Chakrabarty RK, Moosmüller H, Arnott WP et al (2007) Light scattering and absorption by fractal-like carbonaceous chain aggregates: comparison of theories and experiment. Appl Opt 46:6990–7006

Chakrabarty RK, Moosmüller H, Arnott WP et al (2009) Low fractal dimension cluster-dilute soot aggregates from a premixed flame. Phys Rev Lett 102:235504

Chakrabarty RK, Arnold IJ, Francisco DM et al (2013) Black and brown carbon fractal aggregates from combustion of two fuels widely used in Asian rituals. J Quant Spectrosc Radiat 122:25–30

Chakrabarty RK, Beres ND, Moosmüller H et al (2014) Corrigendum: soot super aggregates from flaming wildfires and their direct radiative forcing. Sci Rep 4:5508

Chang H, Charalampopoulos TT (1990) Determination of the wavelength dependence of refractive indices of flame soot. Proc R Soc A 430:577–591

Charalampopoulos TT, Shu G (2002) Effects of polydispersity of chainlike aggregates on light scattering properties and data inversion. Appl Opt 41:723–733

Chen C, Fan X, Shaltout T et al (2016) An unexpected restructuring of combustion soot aggregates by subnanometer coatings of polycyclic aromatic hydrocarbons. Geophys Res Lett 43:11080–11088

Cheng T, Wu Y, Chen H (2014) Effects of morphology on the radiative properties of internally mixed light absorbing carbon aerosols with different aging status. Opt Express 22:15904–15917

China S, Mazzoleni C, Gorkowski K et al (2013) Morphology and mixing state of individual freshly emitted wildfire carbonaceous particles. Nat Commun 4:2122

Chung CE, Ramanathan V, Decremer D (2012a) Observationally constrained estimates of carbonaceous aerosol radiative forcing. Proc Natl Acad Sci USA 109:11624–11629

Chung CE, Lee K, Müller D (2012b) Effect of internal mixture on black carbon radiative forcing. Tellus B 64:1–13

Conant WC, Nenes A, Seinfeld JH (2002) Black carbon radiative heating effects on cloud microphysics and implications for the aerosol indirect effect 1. Extended Köhler theory. J Geophys Res 107. https://doi.org/10.1029/2002jd002094

Coz E, Leck C (2011) Morphology and state of mixture of atmospheric soot aggregates during the winter season over Southern Asia—a quantitative approach. Tellus B 63:107–116

D'Almeida GA, Koepke P, Shettle EP (1991) Atmospheric aerosols: global climatology and radiative characteristics. J Med Microbiol 54:55–61

Dalzell WH, Sarofim AF (1969) Optical constants of soot and their application to heat-flux calculations. J Heat Trans 91:100–104

Dankers S, Leipertz A (2004) Determination of primary particle size distributions from time-resolved laser-induced incandescence measurements. Appl Opt 43:3726–3731

Dlugach JM, Mishchenko MI (2015) Scattering properties of heterogeneous mineral particles with absorbing inclusions. J Quant Spectrosc Radiat 162:89–94

Dobbins RA, Megaridis CM (1991) Absorption and scattering of light by polydisperse aggregates. Appl Opt 30:4747–4754

Doner N, Liu F (2017) Impact of morphology on the radiative properties of fractal soot aggregates. J Quant Spectrosc Radiat 187:10–19

Dong J, Zhao JM, Liu L (2015) Morphological effects on the radiative properties of soot aerosols in different internally mixing states with sulfate. J Quant Spectrosc Radiat 165:43–55

Draine BT, Flatau PJ (1994) Discrete-dipole approximation for scattering calculations. J Opt Soc Am 11:1491–1499

Dukhin AS, Fluck D, Goetz PJ et al (2007) Characterization of fractal particles using acoustics, electroacoustics, light scattering, image analysis, and conductivity. Langmuir 23:5338–5351

Eggersdorfer ML, Pratsinis SE (2012) The structure of agglomerates consisting of polydisperse particles. Aerosol Sci Tech 46:347–353

Farias TL, Köylü ÜÖ, Carvalho MG (1996a) Effects of polydispersity of aggregates and primary particles on radiative properties of simulated soot. J Quant Spectrosc Radiat 55:357–371

Farias TL, Köylü ÜÖ, Carvalho MG (1996b) Range of validity of the Rayleigh-Debye-Gans theory for optics of fractal aggregates. Appl Opt 35:6560–6567

Filippov AV, Zurita M, Rosner DE (2000) Fractal-like aggregates: relation between morphology and physical properties. J Colloid Interface Sci 229:261–273

Freney EJ, Adachi K, Buseck PR (2010) Internally mixed atmospheric aerosol particles: hygroscopic growth and light scattering. J Geophys Res 115. https://doi.org/10.1029/2009jd013558

Fuller KA, Malm WC, Kreidenweis SM (1999) Effects of mixing on extinction by carbonaceous particles. J Geophys Res 104:15941–15954

Ganguly D, Jayaraman A, Gadhavi H et al (2005) Features in wavelength dependence of aerosol absorption observed over central India. Geophys Res Lett 32. https://doi.org/10.1029/2005gl023023

Gao RS, Schwarz JP, Kelly KK et al (2007) A novel method for estimating light-scattering properties of soot aerosols using a modified single-particle soot photometer. Aerosol Sci Tech 41:125–135

Gwaze P, Schmid O, Annegarn HJ et al (2006) Comparison of three methods of fractal analysis applied to soot aggregates from wood combustion. J Aerosol Sci 37:820–838

Gyawali M, Arnott WP, Zaveri RA et al (2012) Photoacoustic optical properties at UV, VIS, and near IR wavelength for laboratory generated and winder time ambient unban aerosols. Atmos Chem Phys 12:2587–2601

Hanel G (1976) The properties of atmospheric aerosol particles as functions of the relative humidity at thermodynamic equilibrium with the surrounding moist air. Adv Geophys 19:73–188

He C, Li Q, Liou KN et al (2016) Microphysics-based black carbon aging in a global CTM: constraints from HIPPO observations and implications for global black carbon budget. Atmos Chem Phys 16:3077–3098

He C, Liou KN, Takano Y et al (2018) Impact of grain shape and multiple black carbon internal mixing on snow albedo: parameterization and radiative effect analysis. J Geophys Res 123. https://doi.org/10.1002/2017jd027752

Herd CR, McDonald GC, Hess WM (1992) Morphology of carbon-black aggregates: fractal versus euclidean geometry. Rubber Chem Technol 65:107–129

Hess M, Koepke P, Schult I (1998) Optical properties of aerosols and clouds: the software package OPAC. B Am Meteorol Soc 79:831–844

IPCC (2013) Climate change 2013: the physical science basis. Cambridge University Press

Jacobson MZ (2001) Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols. Nature 409:695–697

Jullien R, Kolb M (1984) Hierarchical model for chemically limited cluster-cluster aggregation. J Phys A 17:639–643

Kahnert FM (2003) Numerical methods in electromagnetic scattering theory. J Quant Spectrosc Radiat 79–80:775–824

Kahnert M, Devasthale A (2011) Black carbon fractal morphology and short-wave radiative impact: a modelling study. Atmos Chem Phys 11:11745–11759

Kahnert M, Nousiainen T, Lindqvist H et al (2012) Optical properties of light absorbing carbon aggregates mixed with sulfate: assessment of different model geometries for climate forcing calculations. Opt Express 20:10042–10058

Kamimoto T, Shimono M, Kase M (2007) Dynamic measurements of soot aggregate size in diesel exhaust by a light scattering method. J Phys: Conf Ser 85:65–70

Kirchstetter TW, Novakov T (2007) Controlled generation of black carbon particles from a diffusion flame and applications in evaluating black carbon measurement methods. Atmos Environ 41:1874–1888

Kirchstetter TW, Novakov T, Hobbs PV (2004) Evidence that the spectral dependence of light absorption by aerosols is affected by organic carbon. J Geophys Res 109. https://doi.org/10.1029/2004jd004999

Koch D, Del Genio AD (2010) Black carbon semi-direct effects on cloud cover: review and synthesis. Atmos Chem Phys 10:7685–7696

Kolokolova L, Kimura H, Ziegler K et al (2006) Lighter-scattering properties of random-oriented aggregates: do they represent the properties of an ensemble aggregates? J Quant Spectrosc Radiat 100:199–206

Köylü ÜÖ, Faeth GM (1992) Structure of overfire soot in buoyant turbulent-diffusion flames at long residence times. Combust Flame 89:140–156

Köylü ÜÖ, Faeth GM (1994) Optical properties of soot in buoyant laminar diffusion flames. J Heat Trans 116:971–979

Krekov GM (1993) Models of atmospheric aerosols. Aerosol effects on climate 9–72

Lack DA, Cappa CD (2010) Impact of brown and clear carbon on light absorption enhancement, single scatter albedo and absorption wavelength dependence of black carbon. Atmos Chem Phys 10:4207–4220

Lack DA, Langridge JM (2013) On the attribution of black carbon and brown carbon light absorption using the Angstrom exponent. Atmos Chem Phys 13:10535–10543

Lawless PA, Rodes CE, Ensor DS (2004) Multiwavelength absorbance of filter deposits for determination of environmental tobacco smoke and black carbon. Atmos Environ 38:3373–3383

Lee KW (1983) Change of particle size distribution during Brownian coagulation. J Colloid Interface Sci 92:315–325

Lehre T, Jungfleisch B, Suntz R et al (2003) Size distributions of nanoscaled particles and gas temperatures from time-resolved laser-induced-incandescence measurements. Appl Opt 42:2021–2030

Lesins G, Chylek P, Lohmann U (2002) A study of internal and external mixing scenarios and its effect on aerosol optical properties and direct radiative forcing. J Geophys Res 107. https://doi.org/10.1029/2001jd00973

Li J, Anderson JR, Buseck PR (2003) TEM study of aerosol particles from clean and polluted marine boundary layers over the North Atlantic. J Geophys Res 108. https://doi.org/10.1029/2002jd002106

Li H, Liu C, Bi L et al (2010) Numerical accuracy of "equivalent" spherical approximations for computing ensemble-averaged scattering properties of fractal soot aggregates. J Quant Spectrosc Radiat 111:2127–2132

Li J, Liu C, Yin Y et al (2016) Numerical investigation on the Ångström exponent of black carbon aerosol. J Geophys Res 121:3506–3518

Liou KN (2002) An introduction to atmospheric radiation

Liou KN, Takano Y, Yang P (2011) Light absorption and scattering by aggregates: application to black carbon and snow grains. J Quant Spectrosc Radiat 112:1581–1594

Liu QH (1997) The PSTD algorithm: a time-domain method requiring only two cells per wavelength. Microw Opt Tech Lett 15:158–165

Liu L, Mishchenko MI (2005) Effects of aggregation on scattering and radiative properties of soot aerosols. J Geophys Res 110. https://doi.org/10.1029/2004jd005649

Liu L, Mishchenko MI (2007) Scattering and radiative properties of complex soot and soot-containing aggregate particles. J Quant Spectrosc Radiat 106:262–273

Liu F, Smallwood GJ (2010) Radiative properties of numerically generated fractal soot aggregates: the importance of configuration averaging. J Heat Trans 132:207–214

Liu F, Smallwood GJ (2011) The effect of particle aggregation on the absorption and emission properties of mono- and poly-disperse soot aggregates. Appl Phys B 104:343–355

Liu F, Stagg BJ, Snelling DR et al (2006) Effects of primary soot particle size distribution on the temperature of soot particles heated by a nanosecond pulsed laser in an atmospheric laminar diffusion flame. Int J Heat Mass Transfer 49:777–788

Liu L, Mishchenko MI, Arnott WP (2008) A study of radiative properties of fractal soot aggregates using the superposition T-matrix method. J Quant Spectrosc Radiat 109:2656–2663

Liu C, Panetta RL, Yang P (2012a) Application of the pseudo-spectral time domain method to compute particle single-scattering properties for size parameters up to 200. J Quant Spectrosc Radiat 113:1728–1740

Liu C, Panetta RL, Yang P (2012b) The influence of water coating on the optical scattering properites of fractal soot aggregates. Aerosol Sci Tech 46:31–43

Liu D, Allan J, Whitehead J et al (2013) Ambient black carbon particle hygroscopic properties controlled by mixing state and composition. Atmos Chem Phys 13:2015–2029

Liu C, Yin Y, Hu F et al (2015a) The effects of monomer size distribution on the radiative properties of black carbon aggregates. Aerosol Sei Tech 49:928–940

Liu D, Taylor JW, Young DE et al (2015b) The effect of complex black carbon microphysics on the determination of the optical properties of brown carbon. Geophys Res Lett 42:613–619

Liu C, Chung CE, Zhang F et al (2016a) The colors of biomass burning aerosols in the atmosphere. Sci Rep 6:28267

Liu F, Yon J, Bescond A (2016b) On the radiative properties of soot aggregates. Part 2: Effects of coating. J Quant Spectrosc Radiat 172:134–145

Liu C, Li J, Yin Y et al (2017a) Optical properties of black carbon aggregates with non-absorptive coating. J Quant Spectrosc Radiat 187:443–452

Liu D, Whitehead J, Alfarra MR et al (2017b) Black carbon absorption enhancement in the atmosphere determined by particle mixing state. Nat Geosci 10:184–188

Liu C, Teng S, Zhu Y, et al (2018) Performance of discretize dipole approximation for optical property simulations of black carbon aggregates. J Quant Spectrosc Radiat 221:98–109

Liu C, Xu X, Yin Y, et al (2019) Black carbon aggregates: an optical property database. J Quant Spectrosc Radiat 222–223:170–179

Lohmann U, Feichter J (2005) Global indirect aerosol effects: a review. Atmos Chem Phys 5:715–737

Lu Z, Streets DG, Winijkul E et al (2015) Light absorption properties and radiative effects of primary organic aerosol emissions. Environ Sci Tech 49:4868–4877

Mackowski DW (1994) Calculation of total cross sections of multiple sphere cluster. J Opt Soc Am A 11:2851–2861

Mackowski DW (2006) A simplified model to predict the effects of aggregation on the absorption properties of soot particles. J Quant Spectrosc Radiat 100:237–249

Mackowski DW (2014) A general superposition solution for electromagnetic scattering by multiple spherical domains of optically active media. J Quant Spectrosc Radiat 133:264–270

Mackowski DW, Mishchenko MI (1996) Calculation of the T-matrix and the scattering matrix for ensembles of spheres. J Opt Soc Am A 13:2266–2278

Mackowski DW, Mishchenko MI (2011) A multiple sphere T-matrix Fortran code for use on parallel computer clusters. J Quant Spectrosc Radiat 112:2182–2192

Martins JV, Artaxo P, Liousse C et al (1998) Effects of black carbon content, particle size, and mixing on light absorption by aerosols from biomass burning in Brazil. J Geophys Res 103:32041–32050

Menon S, Hansen J, Nazarenko L et al (2002) Climate effects of black carbon aerosols in China and India. Science 297:2250–2253

Mie G (1908) Beitrge zur optik trber medien, speziell kolloidaler metallsungen. Ann Phys 25:377–445

Mishchenko MI, Travis LD, Lacis AA (2002) Scattering, absorption, and emission of light by small particles. Cambridge University Press, Cambridge, UK

Mishchenko MI, Geogdzhayev IV, Cairns B et al (2007) Past, present, and future of global aerosol climatologies derived from satellite observations: a perspective. J Quant Spectrosc Radiat 106:325–347

Moffet RC, Prather KA (2009) In-situ measurements of the mixing state and optical properties of soot with implications for radiative forcing estimates. Proc Natl Acad Sci USA 106:11872–11877

Moosmüller H, Arnott WP (2009) Particle optics in the Rayleigh regime. J Air Waste Manag 59:1028–1031

Moosmüller H, Chakrabarty RK, Arnott WP (2009) Aerosol light absorption and its measurement: a review. J Quant Spectrosc Radiat 110:844–878

Moteki N (2016) Discrete dipole approximation for black carbon-containing aerosols in arbitrary mixing state: a hybrid discretization scheme. J Quant Spectrosc Radiat 178:306–314

Moteki N, Kondo Y, Nakamura SI (2010) Method to measure refractive indices of small nonspherical particles: application to black carbon particles. J Aerosol Sci 41:513–521

Oh C, Sorensen CM (1998) Structure factor of diffusion-limited aggregation clusters: local structure and non-self-similarity. Phys Rev E 57:784–790

Painter TH, Bryant AC, Skiles MK (2012) Radiative forcing by light absorbing impurities in snow from modis surface reflectance data. Geophys Res Lett 39:17502

Peng J, Hu M, Guo S et al (2016) Markedly enhanced absorption and direct radiative forcing of black carbon under polluted urban environments. Proc Natl Acad Sci USA 113:4266–4271

Pirjola L, Niemi JV, Saarikoski S et al (2017) Physical and chemical characterization of urban winter-time aerosols by mobile measurements in Helsinki, Finland. Atmos Environ 158:60–75

Posfai M, Gelencser A, Simonics R et al (2004) Atmospheric tar balls: particles from biomass and biofuel burning. J Geophys Res 109. https://doi.org/10.1029/2003jd004169

Querry MR (1987) Optical constants of minerals and other materials from the millimeter to the ultraviolet, CRDEC-CR-88009. Aberdeen Proving Ground, Maryland

Radney JG, You R, Ma X et al (2014) Dependence of soot optical properties on particle morphology: measurements and model comparisons. Environ Sci Tech 48:3169–3176

Ramanathan V, Carmichael G (2008) Global and reginal climate change due to black carbon. Nat Geosci 1:221–227

Reddington CL, McMeeking G, Mann GW et al (2013) The mass and number size distribution of black carbon aerosol over Europe. Atmos Chem Phys 13:4917–4939

Riemer N, West M, Zaveri R et al (2010) Estimating black carbon aging time-scales with a particle-resolved aerosol model. J Aerosol Sci 41:143–158

Russell PB, Bergstrom RW, Shinozuka Y et al (2010) Absorption Ångström exponent in AERONET and related data as an indicator of aerosol composition. Atmos Chem Phys 10:1155–1169

Sato M, Hansen J, Koch D et al (2003) Global atmospheric black carbon inferred from aeronet. Proc Natl Acad Sci USA 100:6319–6324

Scarnato BV, Vahidinia S, Richard DT et al (2013) Effects of internal mixing and aggregate morphology on optical properties of black carbon using a discrete dipole approximation model. Atmos Chem Phys 13:5089–5101

Schmid O, Artaxo P et al (2006) Spectral light absorption by ambient aerosols influenced by biomass burning in the Amazon Basin. I: Comparison and field calibration of absorption measurement techniques. Atmos Chem Phys 6:3443–3462

Schnaiter M, Linke C, Möhler O et al (2005) Absorption amplification of black carbon internally mixed with secondary organic aerosol. J Geophys Res 110. https://doi.org/10.1029/2005jd006046

Schwarz JP, Gao RS, Spackman JR et al (2008) Measurement of the mixing state, mass, and optical size of individual black carbon particles in urban and biomass burning emissions. Geophys Res Lett 35:L13810

Sharma S, Ishizawa M, Chan D et al (2013) 16-year simulation of arctic black carbon: transport, source contribution, and sensitivity analysis on deposition. J Geophys Res 118:943–964

Shiraiwa M, Kondo Y, Iwamoto T et al (2010) Amplification of light absorption of black carbon by organic coating. Aerosol Sci Tech 44:46–54

Shrestha G, Traina SJ, Swanston CW (2010) Black carbon's properties and role in the environment: a comprehensive review. Sustainability 2:294–320

Skorupski K, Mroczka J (2014) Effect of the necking phenomenon on the optical properties of soot particles. J Quant Spectrosc Radiat 141:40–48

Smith AJA, Peters DM, Mcpheat R et al (2015) Measuring black carbon spectral extinction in the visible and infrared. J Geophys Res 120:9670–9683

Sorensen CM (2001) Light scattering by fractal aggregates: a review. Aerosol Sci Tech 35:648–687

Sorensen CM (2011) The mobility of fractal aggregates: a review. Aerosol Sci Tech 45:765–779

Sorensen CM, Roberts GC (1997) The prefactor of fractal aggregates. J Colloid Interface Sci 186:447–452

Sorensen CM, Cai J, Lu N (1992) Test of static structure factors for describing light scattering from fractal soot aggregates. Langmuir 8:2064–2069

Sorensen CM, Yon J, Liu F et al (2018) Light scattering and absorption by fractal aggregate including soot. J Quant Spectrosc Radiat 217:459–473

Stagg BJ, Charalampopoulos TT (1993) Refractive indices of pyrolytic graphite, amorphous carbon, and flame soot in the temperature range 25–600 °C. Combust Flame 94:381–396

Subramanian R, Roden CA et al (2007) Yellow beads and missing particles: trouble ahead for filter-based absorption measurements. Aerosol Sci Tech 41:630–637

Tumolva L, Park JY, Kim JS et al (2010) Morphological and elemental classification of freshly emitted soot particles and atmospheric ultrafine particles using the TEM/EDS. Aerosol Sci Tech 44:202–215

Van-Hulle P, Weill ME, Talbaut M et al (2002) Comparison of numerical studies characterizing optical properties of soot aggregates for improved EXSCA measurements. Part Part Syst Char 19:47–57

Wang QY, Huang RJ, Cao JJ et al (2015) Black carbon aerosol in winter northeastern Qinghai-Tibetan Plateau, China: the source, mixing state and optical properties. Atmos Chem Phys 15:13059–13069

Wang Y, Liu F, He C et al (2017) Fractal dimensions and mixing structures of soot particles during atmospheric processing. Environ Sci Tech Lett 4:487–493

Witter TA, Sander LM (1981) Diffusion-limited aggregation, a kinetic critical phenomenon. Phys Rev Lett 47:1400–1403

Wu Y, Cheng T, Zheng L et al (2014) A Study of optical properties of soot aggregates composed of poly-disperse monomers using the superposition T-matrix method. Aerosol Sci Tech 49:941–949

Wu Y, Cheng T, Zheng L et al (2016) Black carbon radiative forcing at TOA decreased during aging. Sci Rep 6:38592

Xu YL (1995) Electromagnetic scattering by an aggregate of spheres. Appl Opt 34:4573–4588

Xu YL, Gustafson BÅS (2001) A generalized multiparticle Mie-solution: further experimental verification. J Quant Spectrosc Radiat 70:395–419

Xu YL, Khlebtsov NG (2003) Orientation-averaged radiative properties of an arbitrary configuration of scatterers. J Quant Spectrosc Radiat 79:1121–1137

Yang P, Liou KN (1996) Finite-difference time domain method for light scattering by small ice crystals in three-dimensional space. J Opt Soc Am A 13:2072–2085

Yon J, Bescond A, Liu F (2015) On the radiative properties of soot aggregates. Part 1: Necking and overlapping. J Quant Spectrosc Radiat 162:197–206

Yurkin MA, Hoekstra AG (2007) The discrete dipole approximation: an overview and recent developments. J Quant Spectrosc Radiat 106:558–589

Yurkin MA, Hoekstra AG (2011) The discrete-dipole-approximation code ADDA: capabilities and known limitations. J Quant Spectrosc Radiat 112:2234–2247

Zhang R, Khalizov AF, Pagels J et al (2008) Variability in morphology, hygroscopicity, and optical properties of soot aerosols during atmospheric processing. Proc Natl Acad Sci USA 105:10291–10296

Index